Game, Set, Maths

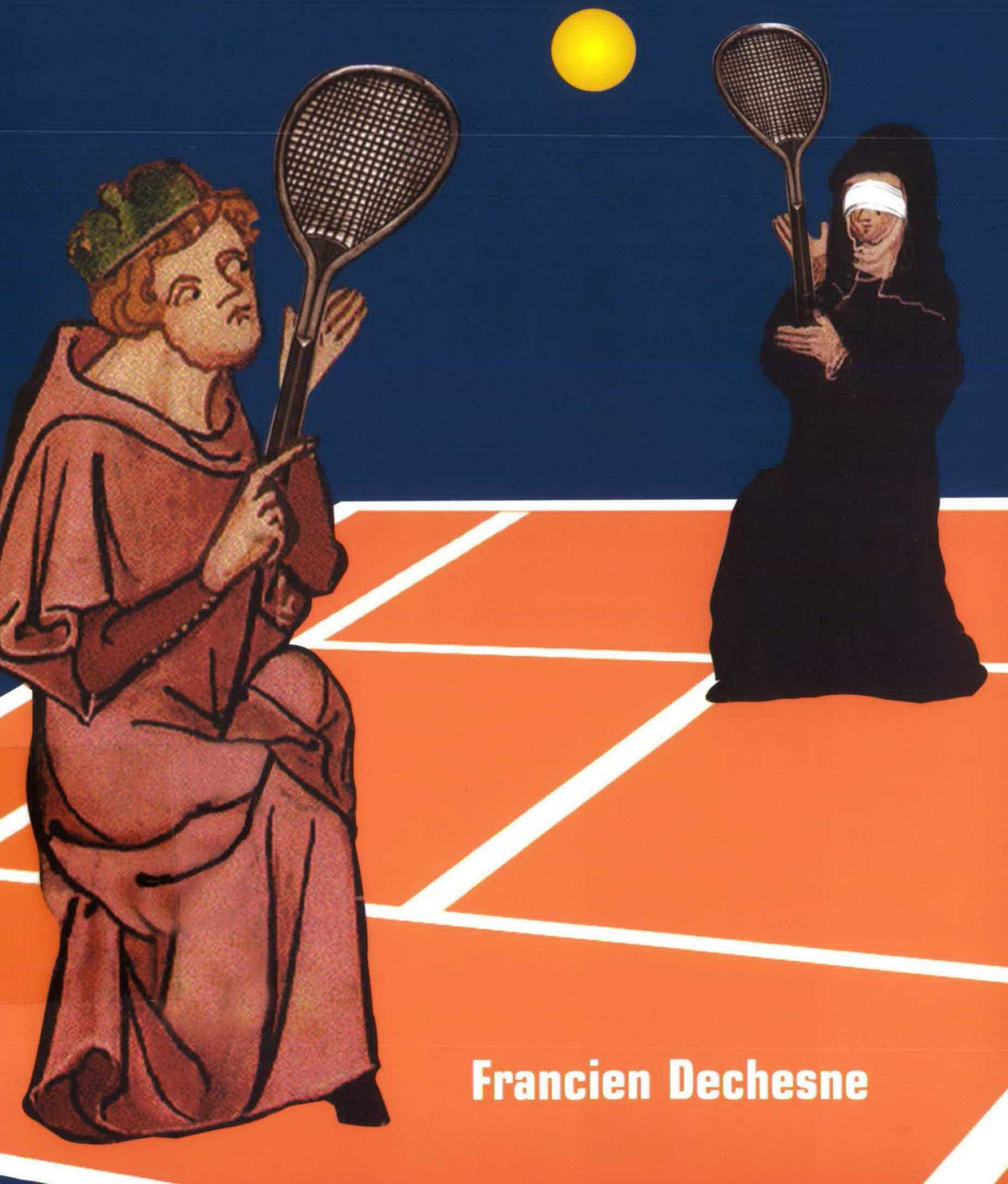
Dechesne, F.

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

Game, Set, Maths:

formal investigations into logic with imperfect information



Francien Dechesne

Game, Set, Maths:
formal investigations into logic with imperfect information

Druk: Printservice, Technische Universiteit Eindhoven

Omslagontwerp: Joost van den Broek.

GAME, SET, MATHS:

FORMAL INVESTIGATIONS INTO LOGIC WITH IMPERFECT INFORMATION

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op maandag 21 maart 2005 om 16:15 uur

door

Francien Dechesne,

geboren op 26 oktober 1971 te Nijmegen.

Promotores:
Prof. Dr. H.C.M. de Swart
Prof. Dr. J.C.M. Baeten

Copromotor:
Dr. R.P. Nederpelt

Dit onderzoek is uitgevoerd aan de faculteit Wijsbegeerte van de Universiteit van Tilburg, en de faculteit Wiskunde & Informatica van de Technische Universiteit Eindhoven. Het is gefinancierd door het Samenwerkingsorgaan Brabantse Universiteiten (SOBU), onder projectcode 99-M, *Hintikka's revolution in the foundations of mathematics.*

Preface

People compare the process of writing a PhD-thesis with all kinds of things. When I had just decided to undertake this adventure, I was very happy to hear from a recent graduate that he had found it to be quite like a long solo journey, probably comparable to the solo bike trip to Italy I had just returned from.

Being a traveller of the type for whom the planning of the route, at home on the map, is almost half of the pleasure of the undertaking, I experienced the scientific journey to be very different in this particular respect. While on the road, a map gives you an indication of your progress and the distance to your target, plus an overview of alternative routes. In science, there turned out to be no map, or only a partial and changing one. An important part of the journey seems to consist of drawing a map, an activity that I had to learn. Fortunately, there were many people to help me find my way to reach my destination.

I thank my supervisors Harrie de Swart and Rob Nederpelt for providing me with the opportunity to make this journey. I am grateful for their continuous support and belief in me, and for their persistent attempts to irradicate my persistent doubts. I also thank my second promotor Jos Baeten, for his interest in my progress and the pleasant environment of his Formal Methods group in Eindhoven as my home away from home: the department of Philosophy in Tilburg.

I specifically want to thank Theo Janssen, for our cooperation, our many discussions, and for his interest in my work and progress. This thesis contains many issues and ideas discussed at our meetings in Amsterdam, and in our joint work. I also thank Xavier Caicedo, whose fruitful ideas contributed substantially to the work presented in chapter 6 of this thesis.

I thank Gabriel Sandu, Johan van Benthem, Theo Janssen, Reinhard Muskens and Elias Thijsse, for their willingness to be in my committee. I value the time they took to critically read the manuscript of this thesis, and appreciate their useful remarks.

Working at two very different departments at two universities, has made me feel like I was always riding in a peloton. Despite the fact that the subject of my research was a bit off-topic for both the philosophers in Tilburg and the computer scientists in Eindhoven, I have always felt at home at both of the *Brabantse Universiteiten*. This is due to the large number of colleagues, all of whom I would like to thank for their company and friendship through the years. As my space here is limited, I can only mention a few of them, my room- and lunchmates over the years: Eliora, Agnieszka, Roland, Mandy, Luigi, Joris, Anton, Michiel, Michael, Ronald, Jan, Timmm, Martijno, Suzana, Georgi, Ana, Uzma, Nikola, Cas, Jasen, Hugo, Michiel, Simona and Bas.

The Dutch Research School in Logic (OzsL) provided me with the opportunity to meet logicians from Groningen, Utrecht, and in particular the ILLC in Amsterdam, on a regular basis. The possibility to discuss my work with these people, and their stimulating suggestions, have been invaluable for my work and motivation. The *schoolweken*, summerschools, conferences and talks at the ILLC were always great if only because they were occasions to meet again. In particular, thanks to Marc Pauly, Barteld Kooi, Paul Harrenstein, Merlijn Sevenster, Boudewijn de Bruin, Clemens Grabmayer, Joost Joosten, Nick Bezhanishvili, Clemens Kupke, Balder ten Cate and Rosja Mastop for being such great company at several occasions.

Let me thank all my friends for the patience of waiting for me to finally call or visit when I promised to (a promise that I did not keep in too many cases). I am looking forward to making up for it in the near future. Other people may have heard a bit too much from me over the last period: thanks to Adriaan, Pieter, Arthur, and Jeroen for all the sense and nonsense we shared on our IRC-channel.

The last stretches of this ride have been a bit like the final kilometers of the Mont Ventoux: you know you're almost there, but after every corner, another hidden turn appears. Joost, thank you very much for designing the colorful cover of this thesis, when I could only see the grey rocks above the tree line. Adriaan and Twan, thank you for turning into pinguins for me and with me at the *cérémonie protocolaire* at the top. Your friendship has been a major support for me along the climb.

I am immensely grateful to my parents, my sister Marieke and my brother Mark, for their love, support and inspiration. Together they form the firm base from which each of my adventures begins, and the safe home that I can always return to.

Tijn, it is magic that you happened to be on the path that I decided to take. Let's take our bike and continue our journey together (back to back, hopefully a bit more laid-back than in the last few months). It will be the ride of my life.

Francien Dechesne
January, 2005

Contents

Chapter 1

Introduction

In his book *The Principles of Mathematics Revisited*, which appeared in 1996, the Finnish philosopher and logician Jaakko Hintikka (1929) presents a 'new and better basic logic' to replace classical first order logic, and claims that this logic can give a essential new impulse to the foundations of mathematics. The book title bears a clear reference to Bertrand Russell's *The Principles of Mathematics* [Rus03], in which he set out the lines of thought that led to the later, more technical, three volume work *Principia Mathematica* with A. N. Whitehead ([WR13]). It may be clear from its aspiration to be *the Principles* for the next century, that Hintikka's book aims to give inspiration by setting out lines of thought, rather than to support the claims by providing mathematically precise accounts for them.

In the book, Hintikka proposes *Independence Friendly logic*, IF-logic for short, to replace first order predicate logic. IF-logic comes with a *game theoretical semantics*, which is to replace the usual Tarskian semantics. This thesis aims to give a mathematically precise account of this logic and its semantics. In the process, our attention is drawn to quite a number of subtleties (e.g. in the syntactic choices). Also, we investigate what we can learn from results from game theory, given that this logic is interpreted by game theoretical semantics.

In this first chapter, we give an informal, general introduction to the subject, and describe a bit of its background.

1.1 Hintikka's view on logic

In [Hin96, Ch. 2], Hintikka distinguishes three functions for logic: logic as a means of expressing (mathematical) propositions ('the descriptive function'), logic as the study of relations of logical consequence (proof theory: 'the deductive function'), and logic as a medium for axiomatic set theory. One could say that first order logic has great merits on all these three functions: it has considerable expressive power (e.g. larger than propositional logic and the Aristotelian syllogisms), sound and complete deductive systems, and it *is* the medium for axiomatic set theory.

In his book, Hintikka argues that the descriptive function is the most important one for the foundations of mathematics. Inference schemes are based on the model theoretic meaning of logical constants ([Hin96, p. 21]): they must be *sound*. Hence, he argues, the descriptive function of logic is more basic than the deductive function.

This gives Hintikka his most important argument against classical first order logic: it is not able to define truth within the language, as established by Tarski's impossibility result of 1933 [Tar33] (which is closely related to Gödel's incompleteness result of 1931 [Göd31], cf. [Hin96, p. 15]). Tarski proved that a truth definition for first order logic can only be formulated in a (second order) metalanguage. This means that the expressive power of first order logic is in an essential sense not strong enough.

Tarski's impossibility result holds more generally for formal languages satisfying certain conditions, among which *compositionality*: the meaning of a complex expression is a function of the meanings of its components. This offers a way out: a formal language with non-compositional semantics may be able to define truth within the language. IF-logic with game theoretical semantics *is* such a non-compositional system: satisfaction is only defined for sentences (closed, possibly complex formulas), the components are only evaluated within the *context* of a sentence. Indeed, due to a back-and-forth translation of IF-logic to existential second order logic, it is possible to define a truth predicate within the language of IF-logic ([Hin96, p. 116], [San98]).

With his focus on descriptive power and banning what he calls 'Tarski's curse', Hintikka sacrifices the deductive function of logic. The proposal of the book can be well understood as being part of the branch of foundational research in mathematics called *Extended Model Theory*, which is part of the study of model theoretic logics. Extended model theory looks for logics –mostly extending first order logic– that are able to *capture* certain mathematical properties (e.g. sets being finite, infinite, countable, uncountable, or functions being continuous, or relations being well-orderings). The aim is to design logics that fit closely to a certain part of mathematical practice, for example in the way the language mirrors the mathematician's talk about the property, or by reflecting the structure of the property. [Bar85] gives a nice introduction in the field.

Using Barwise's terminology, Hintikka's work can be seen as an attack on *the first order thesis*: "a view of logic and mathematics which claims that logic *is* first order logic" [Bar85, p. 3]. This view may have spread among mathematicians and logicians due to the interest and hopes for Hilbert's program, and the success of formalizing parts of mathematics in a first order theory like Zermelo-Fraenkel set theory. Hintikka calls this thesis a "dogma ripe for rejection" ([Hin96, p. viii]). Unlike most of extended model theory, in which *a logic* is designed to reason within a class of structures that have a specific mathematical property, Hintikka seems to want to replace the first order thesis with the thesis that *the logic* is IF-logic.

1.2 Game Theoretical Semantics

To introduce the two ingredients of Hintikka's proposal, we start by looking at an example from mathematics.

The work of Cauchy and Weierstrass [1] gave us the so-called '$\epsilon - \delta$'-definitions, that provide us with means to formally express what we mean when we say "arbitrarily close". In these definitions, Cauchy and Weierstrass are among the first to treat variables not as quantities actively changing, an approach that had led to many controversies. Instead, they use variables as static symbols for any member of a set of possible values. Take for example the definition of continuity of a function, which we can nowadays formulate by a formula from first order logic: a function $f : \mathbb{R} \to \mathbb{R}$ is continuous on its domain if

$$\forall x(\forall \epsilon_{>0})(\exists \delta_{>0})\forall y[(|y-x| < \delta) \to (|f(y) - f(x)| < \epsilon)] \tag{1.1}$$

Weierstrass did not have the language of first order logic to express this like we do. But (paraphrasing [Ste92], as quoted in [Hin96, p. 29]), Weierstrass described the quantifications in terms of a game: first, player Epsilon picks a value for x, and tells player Delta how close he wants the function values to be to $f(x)$ by picking a positive value for ϵ. Then player Delta tells player Epsilon how close the originals need to be to x: knowing x and ϵ as chosen by Epsilon, she picks $\delta > 0$. Player Delta wins this little play if $|f(y) - f(x)| < \epsilon$ for each $y \in (x - \delta, x + \delta)$. The function f is continuous if and only if Delta can win *every* play of this game, in other words: if she has a *winning strategy*.

This game interpretation of quantifiers, which apparently was present in the practice of mathematicians even before Frege formalized quantification in his *Begriffsschrift* ([Fre79]), forms the basis for **Game Theoretical Semantics** (GTS). GTS associates with every first order sentence φ and suitable model a so-called semantic game. This game is played by two players: Eloïse, whose goal it is to show that φ holds in the model (she starts in the role of 'verifier'), and Abélard, whose goal it is to show that φ does *not* hold in the model (he starts in the role of 'falsifier'). The game follows the syntactic structure of φ outside-in, and both quantifiers and connectives prompt moves for one of the players. Moves in the game are either choices for assignments to the variables bound by the quantifiers $\exists$ (move for player in the role of verifier) and $\forall$ (move for the player in the role of falsifier) or choices for one of the two subformulas connected by a connective $\vee$ and $\wedge$ (move for verifier and falsifier respectively). The negation sign $\neg$ does not prompt a move for one of the players, but makes the two players change roles.

[1] In the early nineteenth century, Augustin-Louis Cauchy (1789–1857) was professor at the Ecole Polytechnique in Paris, where he taught courses in Analysis. In his classes he consistently used a higher standard of exactness than was usual at the time, and Laplace urged him to publish his course notes as a book. Even though Gauss and Bolzano had used similar approaches, it were the Cauchy's books that were generally accepted. And we still do accept his approach nowadays: the modern definition of 'convergence' hardly differs from the one Cauchy gave. This may also be due to Karl Weierstrass, who corrected many of the errors and flaws that Cauchy made in his writings. ([vR86, p. 40–41])

A play of the game ends with an atomic formula and an assignment to its free variables. If the atomic formula is satisfied by the assignment in the model, then the player currently in the role of verifier wins, otherwise the player currently in the role of falsifier.

These semantic games allow us to define a satisfaction relation in terms of the existence of *winning strategies* (cf. the continuity example above): φ is true in a given model, if there exists a winning strategy for Eloïse in the semantic game, and φ is false in a given model if there exists a winning strategy for Abélard. (We give a more detailed definition of GTS for first order logic in the next chapter.)

1.3 Independence Friendly logic

One of the improvements of Frege's Begriffsschrift over Aristotelian syllogistic (in which no more than one quantifier occurs in each statement), is that by the use of the dependency relation between quantifiers in a formula we can express relationships between the values of the variables. To illustrate the meaning of this dependence we look again at the definition of continuity (1.1). Here for example, δ may depends on x. If this would not be the case, this would result in a different (more strict) notion of continuity: *uniform* continuity. In that case one chosen δ should work for *all* x in the domain of f.

In first order logic, only a certain type of dependence relations between the quantified variables can occur: scopes of quantifiers are always either nested or non-overlapping. The feature that distinguishes the language of IF-logic from the language of first order logic, is the so-called *slash-operator*. With this operator, we can remove a quantification or a connective from the scope of another quantification. We can thereby create more general patterns of dependency between quantifications than in first order logic. For example, it gives us the possibility to change continuity into uniform continuity by a simple application of the slash operator to the quantification over δ (where we abbreviate the quantifier free part of (1.1) with C):

$$\forall x(\forall \epsilon_{>0})(\exists \delta_{>0})_{/x}\forall y[C(\delta, \epsilon, x, y)] \tag{1.2}$$

Of course, we know that we can also express uniform continuity in first order logic, by changing the order of the quantifiers:

$$(\forall \epsilon_{>0})(\exists \delta_{>0})\forall x\forall y[C(\delta, \epsilon, x, y)] \tag{1.3}$$

An argument for the IF-formulation could be that it shows in a direct way the difference between (ordinary) continuity and uniform continuity. But IF-logic not only offers a different way of formulating first order definable properties. With the more general patterns of dependence between quantifiers in a formula, we *are* able to express properties that were not expressible in first order logic. A case in point is the expressibility of infinity of the domain, which we will demonstrate in chapter 3.

For the interpretation of IF-formulas, we hardly need to adapt the game theoretical semantics for first order logic. We only need to define how we deal with the slash operator.

In game theoretical semantics, the scope of a quantifier corresponds to availability of information. If an existential quantifier $\exists y$ is in the scope of a universal quantifier $\forall x$, then Eloïse knows the value previously chosen by Abélard when she is to choose a value for y. Similarly, if Eloïse has to choose between subformulas $\psi(x)$ and $\theta(x)$ connected by a disjunction in the scope of $\forall x$, she can base her choice on the value of x. Semantic games for first order sentences are therefore games of *perfect information.*

By the effect of the slash operator, which can remove existential quantifications and disjunctions from the scope of previous universal quantifications, the semantic games turn into games of *imperfect information.* If we look at the IF-version (1.2) of uniform continuity, we can interpret the slash at the quantification of δ by stating that player Delta does not *know* the value of x when she has to choose a value for δ. This introduction of imperfect information does not alter the semantic game in terms of its rules. However, it restricts the type of strategy a player can use: if Delta does not know the value of x when she chooses a value for δ, she cannot play a strategy that picks δ as function of x.

Finite depth, two-player win-loss games of perfect information have the property that they are *determined,* in the sense that one of the players has a winning strategy (this is a consequence of the Gale-Stewart theorem [GS53]). In particular, semantic games for first order formulas are determined, a fact which, in logical terms, corresponds with the *law of the excluded middle.* However, finite depth, two-player win-loss games of *imperfect* information no longer have the feature of being determined. It follows that the principle of the excluded middle does not hold for IF-logic. This is already witnessed by a simple IF-sentence like $\forall x \exists y_{/x}[x = y]$, as will be demonstrated when we give a more formal definition of IF-logic in chapter 2.

1.4 Some properties of IF-logic

IF-logic not only extends first order logic, it is also an extension of the theory of Henkin quantifiers. In [Hen61], Henkin introduced a 2-dimensional quantification pattern, the *branching-* or Henkin quantifier:

$$\begin{pmatrix} \forall x & \exists y \\ \forall z & \exists u \end{pmatrix} R(x, y, z, u). \tag{1.4}$$

The meaning of this sentence is *defined* by an existential second order formula, obtained through Skolemization:

$$\exists f \exists g \forall x \forall z R(x, f(x), z, g(z)). \tag{1.5}$$

In [Hen61], this definition is *motivated* by a game theoretical argument. One could say that with his Game Theoretical Semantics, Hintikka takes this *motivation* as his *definition* of the semantics for IF-logic. But at the same time, as we will describe in detail in chapter 3, Hintikka's approach to game theoretical semantics for IF-logic in practice uses a similar Skolemization procedure to obtain *truth conditions* for IF-sentences. For example, the IF-sentence

$$\forall x \forall z \exists y_{/z} \exists u_{/x} R(x, y, z, u) \tag{1.6}$$

has, under Hintikka's approach, exactly the second order sentence (1.5) as its truth condition. (Why we add 'under Hintikka's approach' here, and what this means is the topic of section 3.9 of this thesis.)

Branching quantification, and more generally, partially ordered quantification ([Wal70]), naturally share many properties with IF-logic, as all these logics are based on the idea of allowing more general relations of dependence between quantifiers in first order languages. Henkin quantifiers in which (indexed) connectives occur, are also studied, e.g. in [SV92]. A property they all share, is their expressive power: in all cases, it *equals* that of existential second order logic (Σ^1_1). For partially ordered quantification this was proved independently in [Wal70] and [End70].

Even with this shared expressive power, we could say that IF-logic is more general than partially ordered quantification, because it allows for the most general dependence relations. First, IF-logic also allows (unindexed, normal first order) connectives to be made independent. And second, less visibly: the dependence relations in IF-sentences do not have to be transitive, while "partially ordered" implies "transitive". An example of an IF-sentence in which the dependence relation of the quantifiers is not transitive, is the formula

$$\forall x \exists z \exists y_{/x} [x = y]. \tag{1.7}$$

In this formula, y may depend on z and z may depend on x, while the slash operator indicates that y may *not* depend on x. This example, proposed by Hodges [Hod97a, p. 547], gives rise to interesting comments on logic with imperfect information and Hintikka's presentation of it. The formula will recur many times in this thesis. For example, it is the typical example for a phenomenon called *signaling* (a.o. discussed in section 5.6), and it demonstrates that a certain type of *imperfect recall* occurs in semantic games (section 4.7). In section 3.9, Hintikka's interpretation of this formula is used to demonstrate that his approach to IF-logic was probably more inspired by the Skolemization procedure for partially ordered quantification than by the game theory of the semantics he defined.

Now that we introduced the basic ideas of IF-logic, we sum up some of the important properties mentioned in [Hin96]:

- The law of the excluded middle does not hold (e.g. p. 132, and section 3.8 of this thesis);
- By its treatment of strategies as functions, game theoretical semantics incorporates, and thereby "vindicates", the Axiom of Choice (p. 40);
- IF-logic is not compositional (see, for example, pp. 106-112);

- Every IF-first order sentence can be translated into a (classical) existential second order sentence (Σ^1_1) *and* vice versa (pp. 61-63, and section 3.2 of this thesis);

- (Therefore) a truth predicate can be defined within the language (p. 116);

- We can express a number of mathematical concepts that aren't expressible in first order logic, among which: the infinity of the domain of a model (section 3.2.3 of this thesis); that a certain (first order definable) relation is *not* a well-ordering; that two predicates have the same cardinality;

- The following metalogical theorems hold for IF-logic: the compactness theorem, the separation theorem (in a strengthened form), the Downward Löwenheim-Skolem theorem, and Beth's definability theorem (pp. 59-61);

- The class of valid sentences of IF-logic is not axiomatizable, although the class of inconsistent sentences *is* (pp. 66-68);

- In so-called *Extended* IF-logic, which adds a second, weak *contradictory* negation to the language, the following mathematical concepts can also be expressed: that a certain relation *is* a well-ordering; the principle of mathematical induction; the notion of power set (in a certain sense, cf. [KV89]); the Bolzano-Weierstrass Theorem; continuity in the topological sense; transfinite induction (pp. 188–190).

1.5 Related literature

We higlight a fragment of the literature that has appeared in reaction to Hintikka's proposal. First, there is a number of reviews of the book: by Philippe Kreutz [Kre97], Wilfrid Hodges [Hod97b], Harold Hodes [Hds98], Roy Cook and Stewart Shapiro [CS98], Laurence Goldstein [Gol98], David Corfield [Cor98], Harrie de Swart, Tom Verhoeff and Renske Brands [dSVB99], and the more extensive and detailed one, by Neil Tennant [Ten98].

The claim that IF-logic does not admit of a compositional semantics, has given rise to Hodges' paper [Hod97a]. This paper introduces a compositional semantics for open IF-formulas ([Hod97c] presents the same semantics in a slightly different way). This so-called *"trump semantics"*[2] coincides with game theoretical semantics on IF-sentences (if we take it to be without the extra specification introduced by Hintikka, which we will discuss at the end of chapter 3). However, in [SH01], it is argued that there are different types of compositionality. Hodges' semantics is not of the strong type meant in Hintikka's claim.

Caicedo and Krynicki [CK99] give a very similar semantics for open IF-formulas, and use it to prove a prenex normal form theorem for IF-formulas. We will

[2] The name derives from the term 'trump' that is used to indicate a set of partial assignments for which Eloïse has a winning strategy. The word 'trump' is a corruption of 'triumph' ([Hod97a, p. 552]).

extensively come back to this system and the results of the paper in chapters 5 and 6.

An interesting different approach to the issue of compositionality is proposed by Theo M.V. Janssen in [Jan02]. The *Subgame semantics* given in this paper gives a strictly context independent interpretation to open IF-formulas. This makes most formulas containing the slash *undecided*, i.e. neither true nor false. The expressive power of IF-logic with subgame semantics exceeds that of first order logic by the addition of the truth value *undecided*, but it is not clear if it can characterize non-first order properties.

Jouko Väänänen [Vää02] gives a different game semantics for IF-sentences. The moves for Eloïse in his version of semantic games correspond with the choice of functions rather than domain elements. It can be seen as a sort of 'higher order' semantic game, that is played on the Σ^1_1-truth-condition we get in the usual game theoretical semantics. This higher order game is of perfect information, but the role of Abélard is no longer symmetric to that of Eloïse. By definition, Abélard has a winning strategy whenever Eloïse does not have one, so the law of the excluded middle holds in this approach. Väänänen also introduces an IF-Ehrenfeucht-Fraïssé game characterizing IF-definability and IF-elementary equivalence. It is used to define a distributive normal form for IF-logic.

All papers above dealt with IF-logic as extension to classical first order predicate logic. But the idea of generalizing dependence patterns by a slash operator can be applied in all formal languages in which there is such notion as dependence. In [SP01], independence is introduced in propositional logic, and combined with partiality. In [Bra00], the ideas of IF-logic are applied in a modal logic setting, to examine the associated fixpoint logics. Tero Tulenheimo has studied IF-modal logic and its expressive power ([Tul02], [Tul03]). At the end of [San01] an attempt is made towards IF-linear logic.

In [Hin02b], Hintikka suggests that quantum logic could be viewed as a fragment of extended IF-logic. In [San97], Sandu studies the properties of IF-logic in finite models. Finally, Parikh and Väänänen studied the idea of using a slash operator in a dual manner: not to express what information is *not* available (*in*dependence), but to make explicit which information *is* available (dependence). The result is called *Finite Information logic* (FI-logic), and it is shown to be a decidable sublogic of first order logic ([PV03]).

IF-logic features regularly in van Benthem's more general program of exploring and exploiting the interplay between logic and game theory, cf. [vB00b], [vB01], [vB02b], [vB03], [vB04], [vB05].

1.6 Overview of this thesis

As we hope to convey with the first part of the title of this thesis, "*Game, Set, Maths*", especially by the '*Maths*' in it, the aim of our work has been to get a mathematical grip on the concepts introduced in [Hin96]. As mentioned at the start of this chapter, Hintikka did not focus on (formal) details of the concepts he introduces in [Hin96]. One could say that in this thesis, instead of selling the shiny new car with the new features, we open the hood, and study how the motor is constructed from nuts and bolts.

In the next chapter, "Preliminary Definitions", we collect a series of basic definitions, to which we will later refer back. Because Hintikka defines his IF-language in terms of classical first order formulas, we start by giving a definition of the language of first order logic. We then give the definitions of the IF-language $\mathcal{L}_{IF}$ and game theoretical semantics as given by Hintikka, where necessary adapted to our conventions for this thesis. This chapter is meant to be used as reference in the later chapters.

Chapter 3, "Skolemization and Falsity conditions", gives a precise account of the translation procedure from IF-logic to existential second order logic (Σ^1_1) and back. Hintikka's treatment only takes aspects of truth into account, while, by the failure of the law of the excluded middle, falsity has become a second dimension to the descriptive power of an IF-sentence. We show how to formulate falsity conditions, and study the resulting 'two-dimensional' expressive power of IF-logic.

In chapter 4, "Game Theory as formal framework", we take the "Game Theoretical" of GTS seriously (corresponding to the '*Game*' in the title of this thesis). We define a more general language $\mathcal{L}_{IFG}$, and formalize semantic games for the IFG-sentences in game-theoretic terms. We use the so-called extensive form to model them, and look for new insights in the logic by looking at it from a game-theoretic perspective: we study correspondences between logical equivalence and the Thompson transformations, and we study the character of *imperfect recall* as present in IF-semantic games.

In the work of Hintikka and in the previous chapters, only sentences (closed formulas) are interpreted. In chapter 5, "Satisfaction for open formulas", we introduce and comment the semantics for IFG-formulas with free variables. These work with *sets* of valuations in order to be able to interpret independence of free variables (this part therefore corresponds to the '*Set*' in the title). In chapter 6 we then study the prenex normal form theorem of [CK99] that uses this semantics. We highlight several flaws in the original formulations and proofs of the lemmas, and give improved versions of them. It turns out that the main result holds for the natural class of *regular* formulas. We also show that a similar restriction should be made for the claim that IF-logic is a conservative extension of first order logic.

In the last chapter, chapter 7, we end with some general conclusions. We finish with a list of interesting open problems, that were unfortunately out of the scope of this project, and may serve as inspiration for further research.

Chapter 2

Preliminary definitions

In this chapter, we collect precise definitions known from the literature, of basic languages and notions that will be used throughout this thesis: first order logic, game theoretical semantics and IF-logic as proposed by [Hin96], and the language of existential second order logic: Σ_1^1.

The definitions of *game theoretical semantics* and *IF-logic* are the central ones. But we start with definitions of the first order language –because the IF-language is built from it– and Tarski style semantics –because game theoretical semantics is compared with it. We end with a definition of Σ_1^1, as preparation for the next chapter, where the relation between IF-logic and Σ_1^1 is elaborated.

2.1 The first order language $\mathcal{L}_{FOL}$

We define what we will mean when we say '(classical) first order language'. (The definitions in the first sections of this chapter are based on [Sch67], [dS93], and [Fit96]. Because we assume the reader to be familiar to first order logic, we state the definitions without much explanation.)

Definition 2.1.1 (first order signature) *A first order signature is a 3-tuple*

$$\sigma = \langle \mathtt{C}, \mathtt{P}, \mathtt{F} \rangle$$

where[1]

- $\mathtt{C}$ *is a finite or countable set of constant symbols,*
- $\mathtt{P}$ *is a finite or countable set of predicate symbols,* $P^{(m)}$, *where* $m \in \mathbb{N}$ *indicates the fixed arity of the symbol,*
- $\mathtt{F}$ *is a finite or countable set of function symbols* $f^{(n)}$, *where* $n \in \mathbb{N}$ *indicates the fixed arity of the symbol.*

[1] It would be formally more correct to define $\mathtt{P}$ and $\mathtt{F}$ to be arbitrary sets of symbols, and let the signature σ provide 'arity assigning' functions $a_{\mathtt{P}}$ and $a_{\mathtt{F}}$. I chose not to do so, because that level of formality does not seem necessary, and could harm the readability.

Convention: we omit the superscript of any of the symbols in P *or* F, *whenever the arities are clear from the context.*

Definition 2.1.2 (alphabet) *Let σ be a first order signature. The first order language $\mathcal{L}^{\sigma}_{FOL}$ has the following alphabet:*

- *logical symbols: $\neg$ (negation), $\wedge, \vee$ (binary connectives) and $\forall, \exists$ (quantifiers);*
- *punctuation: (,),[,] (parentheses);*
- *countably many variables: $x, y, z, s, t, u, x_1, x_2, x_3, \dots, y_1, y_2, y_3 \dots$;*
- *individual constants: the elements of* C*:*
- *predicate symbols: the elements of* P
- *function symbols: the elements of* F.

Definition 2.1.3 (terms) *The set of terms of the language $\mathcal{L}^{\sigma}_{FOL}$ is defined by:*

- *every variable and every individual constant is a term;*
- *if $f^{(n)} \in \mathtt{F}$, and if $t_1, \dots, t_n$ are terms, then $f(t_1, \dots, t_n)$ is a term;*
- *that's all.*

Definition 2.1.4 (first order formulas) *The set of formulas of the language $\mathcal{L}^{\sigma}_{FOL}$ is defined by:*

- *If $P^{(m)} \in \mathtt{P}$, and if $t_1, \dots, t_m$ are terms, then $P(t_1, \dots, t_m)$ is a formula. These formulas are called* **atomic**.
- *If φ is a formula, then so is $\neg(\varphi)$;*
- *if φ and ψ are formulas, then so are $(\varphi) \vee (\psi)$ and $(\varphi) \wedge (\psi)$;*
- *if φ is a formula, and x a variable, then $\forall x[\varphi]$ and $\exists x[\varphi]$ are also formulas.*
- *That's all.*

We say that φ is a **first order formula** *($\varphi \in \mathcal{L}_{FOL}$) if $\varphi \in \mathcal{L}^{\sigma}_{FOL}$ for some first order signature σ.*

The parentheses are used to avoid ambiguities. The priority order of the connectives is $\neg, \wedge, \vee$. If there is no risk for ambiguities, we will almost always omit the parentheses.

Furthermore, if strict formality is not required, we will not explicitly mention σ, but use predicate symbols $P, R, \dots$ and function symbols $f, g, h, \dots$, and let their arities be clear from the context.

Definition 2.1.5 (subformulas) *The set of subformulas of a first order formula φ is defined inductively:*

- *if* φ *is atomic, then* $Sub(\varphi) = \{\varphi\}$;
- *if* φ *is* $\psi_1 \wedge \psi_2$ *or* $\psi_1 \vee \psi_2$, *then* $Sub(\varphi) = Sub(\psi_1) \cup Sub(\psi_2) \cup \{\varphi\}$;
- *if* φ *is* $\neg\psi, \forall x\psi$ *or* $\exists x\psi$, *then* $Sub(\varphi) = Sub(\psi) \cup \{\varphi\}$.

Definition 2.1.6 (free and bound variables) *The set of free variables and the set of bound variables of a first order formula* φ *are both defined inductively:*

- *if* φ *is atomic, then* $Fv(\varphi)$ *is the set of all variables that occur in* φ, *and* $Bv(\varphi) = \emptyset$,
- *if* φ *is* $\neg\psi$, *then* $Fv(\varphi) = Fv(\psi)$ *and* $Bv(\varphi) = Bv(\psi)$,
- *if* φ *is* $\psi_1 \wedge \psi_2$ *or* $\psi_1 \vee \psi_2$, *then* $Fv(\varphi) = Fv(\psi_1) \cup Fv(\psi_2)$, $Bv(\varphi) = Bv(\psi_1) \cup Bv(\psi_2)$,
- *if* φ *is* $\forall x\psi$ *or* $\exists x\psi$, *then* $Fv(\varphi) = Fv(\psi) \setminus \{x\}$, $Bv(\varphi) = Bv(\psi) \cup \{x\}$.

A formula φ *with* $Fv(\varphi) \neq \emptyset$ *is called an* **open formula**. *A* **sentence** *is a formula* φ *with no free variables:* $Fv(\varphi) = \emptyset$.

Definition 2.1.7 (negation normal form) *A formula* φ *is in* **negation normal form**, *if negation occurs only at atomic level. (I.e.: if* φ *is a formula in negation normal form, and* $\neg\psi$ *is a subformula of* φ, *then* ψ *is atomic).*

Definition 2.1.8 (scope) *If* $Qx[\psi]$ *with* $Q \in \{\forall, \exists\}$ *is a subformula of* φ, *then* ψ *is called the* **scope** *of the quantification* Qx *in* φ.

2.2 First order models

Definition 2.2.1 (first order model) *Let* $\sigma = \langle \mathtt{C}, \mathtt{P}, \mathtt{F} \rangle$ *be a first order signature. A first order model of signature* σ *is a tuple:*

$$\mathfrak{A} = \langle A, I_\mathtt{C}, I_\mathtt{P}, I_\mathtt{F} \rangle$$

where

- A *is a* non-empty *set, and is called the* domain *of* $\mathfrak{A}$
- $I_\mathtt{C} : \mathtt{C} \to A$; *for every constant symbol* $c \in \mathtt{C}$, *we call* $I_\mathtt{C}(c)$ *the interpretation of* c *in* $\mathfrak{A}$.
- $I_\mathtt{P} : \mathtt{P} \to \bigcup_m \mathcal{P}(A^m)$ *with for every predicate symbol* $P^{(m)} \in \mathtt{P}$: $I_\mathtt{P}(P) \subseteq A^m$ *("the interpretation of* $P^{(m)}$ *in* $\mathfrak{A}$*").*
- $I_\mathtt{F} : \mathtt{F} \to \bigcup_n A^{(A^n)}$ *with for every function symbol* $f^{(n)} \in \mathtt{F}$: $I_\mathtt{F}(f) \in A^{(A^n)}$ *("the interpretation of* $f^{(n)}$ *in* $\mathfrak{A}$*").*

As hinted in [Bar85, p. 5], there is a slight disagreement among the adherents of the *first order thesis*, i.e. the view that logic is what is implicit in the logical constants (quantifiers, connectives), as to whether equality (identity, =) should be counted as a logical constant. We choose not to regard '=' to be a logical constant, but it will be present in most models we encounter in this thesis. If it is, we assume its interpretation satisfies the equality axioms. (As we will see in section 3.2.2 in the next chapter, equality *is* needed in Hintikka's approach to IF-logic.)

When discussing an arbitrary model $\mathfrak{A}$, we will use the symbol A to indicate its domain (as in the definition above). Informally, if the sets of symbols of σ contain only a few elements, we will indicate models by writing the interpretations explicitly. For example, if $\sigma = \langle\{c\}, \emptyset, \{f^{(1)}\}\rangle$, then we mean by $\langle \mathbb{N}, 0, S\rangle$ the first order model of signature σ with the natural numbers as domain, the natural number 0 as the interpretation of the constant symbol c, and the successor function S as interpretation for the unary function symbol f.

We will usually omit explicit reference to the first order signatures σ. When we say φ is a first order formula (or $\varphi \in \mathcal{L}_{FOL}$), we mean that $\varphi \in \mathcal{L}^{\sigma}_{FOL}$ for some first order signature σ. Signatures play a role implicitly in the notion of 'suitable' model:

Definition 2.2.2 (suitable model) *If φ is a first order formula, a first order model of signature σ is called* **suitable** *for φ, if $\varphi \in \mathcal{L}^{\sigma}_{FOL}$ (i.e. the model has an interpretation for all symbols in φ; if φ is clear from the context, we will simply call $\mathfrak{A}$ 'a suitable model').*

2.3 Tarski style semantics for first order logic

The classical way of evaluating first order formulas in first order models is by a recursive *satisfaction relation*, which was introduced by Alfred Tarski in 1933. Before we can define it, we need some extra definitions.

We deviate from the usual definition of satisfaction in one respect: we work with *valuations* instead of *assignments*. If we let $\mathtt{Var}$ denote the set of all variables in the language, assignments are functions $v : \mathtt{Var} \to A$ assigning a domain element to *all* variables in the language. The evaluation of first order formulas only depends on the values assigned to the free variables in this formula. However, as will be demonstrated in chapters 5 and 6 of this thesis, for open formulas with the independence operator of IF-logic, the evaluation can be influenced by values that are assigned to variables that do not occur in the formula. Such formulas are therefore evaluated using *partial* assignments, which we call *valuations*.

In the light of a comparison between classical satisfaction and satisfaction for open formulas as part an IF-language (section 5.5), we also use valuations to define satisfaction for first order formulas:

Definition 2.3.1 (valuations) *Let $X \subseteq \mathtt{Var}$ be a set of variables and $\mathfrak{A}$ a first order model. Then a function $v : X \to A$ is called* **a valuation** *(of X in $\mathfrak{A}$). If φ is a first order formula, we call a valuation v* **suitable** *for φ, if $Fv(\varphi) \subseteq dom(v)$ (again, if φ is clear from the context, we will simply call v 'a suitable valuation').*

In practice, all valuations we encounter will be partial assignments with a finite domain. For these we introduce the following explicit notation:

Notation 2.3.2 (explicit notation for valuations) *We write* $v \in A^{\{x_1,\dots,x_k\}}$ *as* $(x_1 \dots x_k \colon a_1 \dots a_k)$, *where* $a_i = v(x_i)$. *For example,* $(xy\colon 01)$ *is the valuation in* $\{0,1\}^{\{x,y\}}$ *assigning the value 0 to* x *and the value 1 to* y.

For the atomic case of the definition of satisfaction, we use the following definition:

Definition 2.3.3 (valuations extended to terms) *Given a first order model* $\mathfrak{A} = \langle A, I_{\mathtt{C}}, I_{\mathtt{P}}, I_{\mathtt{F}} \rangle$ *of signature* σ, *and a set of variables* X, *let* $T_{\sigma,X}$ *denote the set of terms of* $\mathcal{L}^{\sigma}_{FOL}$ *that contain only variables from* X. *A valuation* $v : X \to A$ *is uniquely extended to a function* $\hat{v} : T_{\sigma,X} \to A$ *interpreting the terms, by the following inductive definition:*

- *for every* $x \in X$: $\hat{v}(x) := v(x)$
- *for every individual constant* $c \in \mathtt{C}$: $\hat{v}(c) := I_{\mathtt{C}}(c)$
- *if* $f^{(n)} \in \mathtt{F}$ *and* $t_1, \dots, t_n$ *are terms, then*

$$\hat{v}(f(t_1,\dots,t_n)) := I_{\mathtt{F}}(f)(\hat{v}(t_1),\dots,\hat{v}(t_n)).$$

For the quantifier case, we will need the following definition:

Definition 2.3.4 (x-variants) *If* $v \in A^X$ *is a valuation of* X *in* A, x *an arbitrary variable and* $a \in A$, *then we use the notation* $v(x\colon a)$ *for the valuation* $v' : X \cup \{x\} \to A$, *defined by* $v'(x) := a$ *and* $v'(y) = v(y)$ *for all* $y \in X - \{x\}$. *We call a valuation of the form* $v(x\colon a)$ *an* x-**variant** *of* v.

We can now define satisfaction for first order formulas in the Tarski-style. (The notation we use, with the valuation between brackets rather than before the turnstyle, is chosen to accord with the notation used for satisfaction for open formulas in an IF-language by [CK99], given here in chapter 5).

Definition 2.3.5 (Tarski-style satisfaction) *Let* φ *be a first order formula,* $\mathfrak{A} = \langle A, I_{\mathtt{C}}, I_{\mathtt{P}}, I_{\mathtt{F}} \rangle$ *a suitable model, and* $v : X \to A$ *a suitable valuation. We then define the* satisfaction relation

$$\mathfrak{A} \models \varphi[v]$$

inductively by distinction of the following cases:

(At) If φ *is atomic, say* $\varphi = P(t_1,\dots,t_n)$, *then* $\mathfrak{A} \models \varphi[v]$ *if and only if*

$$(\hat{v}(t_1), \dots \hat{v}(t_n)) \in I_{\mathtt{P}}(P)$$

where $\hat{v}$ *is defined as in definition 2.3.3.*

($\neg$) If $\varphi = \neg\psi$, *then* $\mathfrak{A} \models \varphi[v]$ *if and only if* $\mathfrak{A} \not\models \psi[v]$ *(i.e. not:* $\mathfrak{A} \models \psi[v]$*).*

($\vee$) If $\varphi = \psi_1 \wedge \psi_2$, *then* $\mathfrak{A} \models \varphi[v]$ *if and only if* $\mathfrak{A} \models \psi_1[v]$ *or* $\mathfrak{A} \models \psi_2[v]$.

($\wedge$) *If* $\varphi = \psi_1 \wedge \psi_2$, *then* $\mathfrak{A} \models \varphi[v]$ *if and only if* $\mathfrak{A} \models \psi_1[v]$ *and* $\mathfrak{A} \models \psi_2[v]$.

($\exists$) *If* $\varphi = \exists\psi$, *then* $\mathfrak{A} \models \varphi[v]$ *if and only if there exists an* $a \in A$, *such that* $\mathfrak{A} \models \psi[v(x\colon a)]$.

($\forall$) *If* $\varphi = \exists\psi$, *then* $\mathfrak{A} \models \varphi[v]$ *if and only if for all* $a \in A$, $\mathfrak{A} \models \psi[v(x\colon a)]$.

If $\mathfrak{A} \models \varphi[v]$, *we say that* φ *is* **satisfied** *in* $\mathfrak{A}$ *with respect to* v. *If* φ *is a sentence, we write* $\mathfrak{A} \models \varphi$ *instead of* $\mathfrak{A} \models \varphi[\lambda]$, *where* $\lambda : \emptyset \to A$ *is the valuation with empty domain; if* $\mathfrak{A} \models \varphi$ *we say that* φ *is* **true** *in* $\mathfrak{A}$.

2.4 Game theoretical semantics (GTS)

Game theoretical semantics (GTS) constitutes a notion of satisfaction through the analysis of what we will call *semantic games*. The following definition of semantic games is a slightly adapted version of the definition given by Hintikka in [Hin96, p. 25].

Hintikka's definition does include a rule for negation (role-switch), but there is no component in the definition that keeps track of the role distribution of the two players. Admittedly, for formulas in negation normal form, with which Hintikka usually works, role switches do not really occur. But we prefer to define the semantics for the general case (as Hintikka does as well), and take the negation rule seriously. Therefore, we define the game for two players with names that do not include their roles (Eloïse, Abélard)[2] and take the names *Verifier* and *Falsifier*, used by Hintikka, as names of the roles these players can have in the game. These roles can be seen in analogy to playing White or Black in a game of chess, and negation is like turning the board such that the player who played White now plays Black and vice versa. (Karpov does not *become* Kasparov and vice versa, they only switch roles.)

Definition 2.4.1 (semantic games) *Let* φ *be a first order sentence, and* $\mathfrak{A}$ *a suitable model. The* semantic game $G_{\mathfrak{A}}(\varphi)$ *is played by two players: Abélard and Eloïse. There are two roles they can play in the game:* falsifier *and* verifier. *Each position in the game is described by a triple* $\langle\psi, v, p\rangle$, *where* ψ *is a subformula of* φ, v *a valuation for* ψ *in* $\mathfrak{A}$, *and* $p \in \{-1, +1\}$ *a parameter indicating the role distribution:* $p = 1$ *designates that Abélard plays the role of* falsifier, *and Eloïse plays the role of* verifier, $p = -1$ *indicates reversal of these roles. The game starts from the initial position* $\langle\varphi, \lambda, 1\rangle$. *A play of the game proceeds along the following rules:*

($\exists$) *In a position* $\langle\exists x\psi, v, p\rangle$, *the player in the role of verifier chooses an element* $a \in Dom(\mathfrak{A})$. *The game continues from position* $\langle\psi, v(x\colon a), p\rangle$.

[2] The names Abélard and Eloïse are commonly used for players of logic games, obviously because of the association by their initials of Abélard with the universal quantification $\forall$ and Eloïse with the existential quantification $\exists$. Also, to have one male and one female player has the practical advantage that they can unambiguously be referred to by male and female pronouns respectively. In both respects, e.g. the names Adam and Eve would serve just as well, but we like the fact that the actual Abélard studied (Aristotelian) logic for a while, and Eloïse, as his pupil, may also have known a bit about logic too. See Appendix A for a brief description of their lives.

($\forall$) *In a position* $\langle \forall x\psi, v, p\rangle$, *the player in the role of falsifier chooses an element* $a \in Dom(\mathfrak{A})$. *The game continues from position* $\langle \psi, v(x\colon a), p\rangle$.

($\vee$) *In a position* $\langle \psi_1 \vee \psi_2, v, p\rangle$, *the player in the role of verifier chooses* $i \in \{1, 2\}$. *The game continues from position* $\langle \psi_i, v, p\rangle$.

($\wedge$) *In a position* $\langle \psi_1 \wedge \psi_2, v, p\rangle$, *the player in the role of falsifier chooses* $i \in \{1, 2\}$. *The game continues from position* $\langle \psi_i, v, p\rangle$.

($\neg$) *In a position* $\langle \neg\psi, v, p\rangle$, *the two players switch roles, and the game continues from position* $\langle \psi, v, -p\rangle$.

(*At*) *In a position* $\langle \psi, v, p\rangle$ *with* ψ *atomic, the player currently in the role of verifier wins if* $\mathfrak{A} \models \varphi[v]$, *and the player in the role of falsifier loses. Otherwise, i.e. if* $\mathfrak{A} \not\models \varphi[v]$, *the player in the role of falsifier wins and the player in the role of verifier loses.*

To give a simple example, consider the sentence $\forall x \neg \forall y[x = y]$ in the model $\mathfrak{A} = \langle \{0, 1\}, =\rangle$. A play of the game starts with the choice by Abélard of a value a_1 for x, followed by a role switch, so Eloïse gets to pick a value a_2 for the universally quantified variable y. We then hit the atomic formula $x = y$: Abélard is currently in the role of verifier, so he wins if $a_1 = a_2$, while Eloïse, currently falsifier, wins if $a_1 \neq a_2$.

While we understand very clearly from the definition of semantic games what the plays of such game are, it is not clear what the outcome of a particular play of the game *means* in logical terms (cf. the table on page 38 of [Hin96]). But the semantics is not defined in terms of outcomes of single plays, but by the existence of a winning *strategy*. In our example, it is clear that in this game Eloïse can win every play: whether Abélard chooses 0 or 1, she can pick 1 and 0 respectively. In other words: she has a winning strategy. It is the existence of a winning strategy that gives us the game-theoretic definition of satisfaction for first order sentences:

Definition 2.4.2 (truth and falsity in GTS) *Let* φ *be a first order sentence and* $\mathfrak{A}$ *a suitable model, then we define:*
$\mathfrak{A} \models\!\!\models_{\mathrm{t}} \varphi$ *("φ is true in $\mathfrak{A}$") if and only if Eloïse has a winning strategy in* $G_{\mathfrak{A}}(\varphi)$.
$\mathfrak{A} \models\!\!\models_{\mathrm{f}} \varphi$ *("φ is false in $\mathfrak{A}$") if and only if Abélard has a winning strategy in* $G_{\mathfrak{A}}(\varphi)$.

What we define to be a *strategy* is therefore central to GTS. Hintikka gives no formal definition of the notion of strategy. He describes what he means by the concept as follows: "In my sense, a strategy for a player is a rule that determines which move that player should make in any possible situation that can come up in the course of a play." In practice (i.e. in his work, like in [Hin96]), Hintikka lets strategies be sets of (generalized) Skolem-functions: for example in the semantic game described above, a strategy for Eloïse would be a unary function $f : A \to A$, such that $\langle \mathfrak{A}, f\rangle \models \forall x[x \neq f(x)]$.

We mention some elementary properties of semantic games for first order sentences: as they are finite-depth two-player win-loss games of perfect information, they are determined in the sense that one of the players has a winning strategy. This follows from the Gale-Stewart theorem.[3] Note that, in win-loss games, it can never be the case that both players have a winning strategy at the same time. So, logically:

$$\mathfrak{A} \models_{\mathrm{t}} \varphi \textbf{ or } \mathfrak{A} \models_{\mathrm{f}} \varphi, \text{ while } \textit{not}\text{: } \mathfrak{A} \models_{\mathrm{t}} \varphi \textbf{ and } \mathfrak{A} \models_{\mathrm{f}} \varphi$$

GTS for first order sentences can be proved to coincide with Tarksi semantics, but how this is proved depends on the chosen definition for the concept of strategy. If we take a strategy for a player to be some *function* prescribing *one* choice in every possible situation of the game where this player has to make a move (*determinate strategies*), then the axiom of choice is needed. An alternative is to take a strategy to be a *relation*, prescribing *non-empty sets* of possible choices (*undeterminate strategies*), in which case it can be proved without the axiom of choice. In this thesis all strategies will be of the functional type thereby incorporating the axiom of choice in GTS (cf. [Hin96, p. 40]).

2.5 IF-logic: semantic games of imperfect information

In [Hin96, p. 52], Hintikka defines the language of IF first order logic essentially as follows:

Definition 2.5.1 ($\mathcal{L}_{IF}$: IF-sentences) *Given a first order signature σ, the language $\mathcal{L}^{\sigma}_{IF}$ is determined by the following conditions:*

- *$\mathcal{L}^{\sigma}_{IF}$ contains all sentences of $\mathcal{L}^{\sigma}_{FOL}$ in negation normal form;*
- *If φ is in $\mathcal{L}^{\sigma}_{IF}$ and if a quantification $\exists y$ occurs in φ within the scope of universal quantifiers among which $\forall x_1, \forall x_2, \ldots, \forall x_n$, then the formula resulting from replacing $\exists y$ by $\exists y_{/\{x_1,x_2,\ldots,x_n\}}$ is also in $\mathcal{L}^{\sigma}_{IF}$*
- *If φ is in $\mathcal{L}^{\sigma}_{IF}$ and if a disjunction $\vee$ occurs in φ within the scope of universal quantifiers among which $\forall x_1, \forall x_2, \ldots, \forall x_n$, then the formula resulting from replacing $\vee$ by $\vee_{/\{x_1,x_2,\ldots,x_n\}}$ is also in $\mathcal{L}^{\sigma}_{IF}$*
- *That's all.*

Convention: for simplicity of notation, we usually write the variables under the slash as a sequence ('$x_1, \ldots, x_n$'), rather than as a set ('$\{x_1, \ldots, x_n\}$'). We identify first order quantifications $\exists y$ with $\exists y_{/\emptyset}$ and ordinary disjunction $\vee$ with $\vee_{/\emptyset}$. We say that φ is an **IF-sentence** *($\varphi \in \mathcal{L}_{IF}$) if $\varphi \in \mathcal{L}^{\sigma}_{IF}$ for some first order signature σ.*

[3] It is also common practice to refer to the earlier theorem from [Zer13], but [SW01] warns that most "modern statements of Zermelo's theorem bear only partial relationship to what Zermelo really did." One difference with our situation and the situation of chess discussed by Zermelo, is the fact that in semantic games, the number of *positions* in the game may be infinite (if the model is infinite), while Zermelo assumes finitely many positions. We therefore refer more safely to the Gale-Stewart theorem.

Remark: We deviate from the syntax introduced by Hintikka by omitting the universal quantifiers under the slash: Hintikka writes $\exists y_{/\forall x_1,\ldots,\forall x_n}$. We see no technical reason why mentioning the quantifiers would be necessary, and think the formulas are more easily readable without the quantifiers. In this respect we follow a.o. [Hod97a] and [CK99]. However, mentioning the quantifiers with the variables would make the characteristic fact explicit that in $\mathcal{L}_{IF}$ existential quantifiers and disjunctions are only slashed for universally quantified variables.

The slash operator in a quantification $\exists y_{/x_1,\ldots,x_n}$ removes the quantification from the *scope* of the universal quantifications $\forall x_1, \ldots, \forall x_n$. In other words, by the addition of $_{/x_1,\ldots,x_n}$ to the quantification $\exists y$, this quantification no longer *depends* on the quantifications $\forall x_1, \ldots, \forall x_n$: it makes it **independent**. The name *Independence Friendly* logic, which is usually abbreviated to IF-logic, is due to the feature that the slash operator allows for –"is friendly towards"– a more general class of dependency patterns: in this language we can introduce independence where this was not possible in first order logic (where scopes are either nested or non-overlapping).

Semantics for the IF-language is in a sense *the same* game theoretical semantics as defined for first order logic. The rules for the games (hence the plays of the games) remain the same, and so does the definition of satisfaction. How then are the applications of the slash operator interpreted? The independence introduced by the slash operator at an existential quantifier or disjunction, is interpreted in terms of the information available for Eloïse when she is to make a move. Independence of a quantification $\forall x$, means that Eloïse does not *know* the value previously assigned to x by Abélard. Semantically, the slash operator turns the semantic games for first order logic into games of *imperfect information.*

This is best illustrated by a simple example. Consider $\forall x \exists y_{/x}[x \neq y]$; note how this IF-sentence is the result of application of the slash operator to the negation normal form of the first order formula $\forall x \neg \forall y[x = y]$, which we used as example after the definition of semantic games in the previous section. In the semantic game for this IF-sentence in the model $\mathfrak{A} = \langle \{0, 1\}, = \rangle$, Abélard chooses a value for x, then Eloïse chooses a value for y, but she now does so without knowing the value chosen by Abélard. Does she still have a winning strategy? The answer is easily seen to be 'no': if she chooses 0, then she loses if Abélard happened to choose 0 as well, and similarly for the choice 1. It is also easy to see that similarly, Abélard does not have a winning strategy either. This simple example already illustrates that **the law of the excluded middle fails** in IF-logic with GTS.

The transition from games of perfect information to games of imperfect information (which also allows for the indeterminacy of the semantic games), is hence not visible in an alteration of the games in terms of the rules, or the possible plays. It is reflected in the 'higher order' concept of *strategy.* In Hintikka's approach with Skolem functions, the slash operator in IF-sentences results in a reduction of arguments for these functions (cf. the first order example in the previous section): for the IF-sentence $\forall x \exists y_{/x}[x \neq y]$, Eloïse would have a winning strategy if there were a 0-ary Skolem-function, or, in other words, a Skolem *constant* f, such that

$\langle \mathfrak{A}, f\rangle \models \forall x[x \neq f]$.

2.6 Existential second order logic: Σ_1^1

In preparation of, in particular, the following chapter, we define the language of existential second order logic, Σ_1^1. The Σ-notation comes from the analogy with set-theoretic hierarchy theory, as can be understood from the following quote from [Wal70, p. 537]: "Using the symbols Σ_n^m and Π_n^m of hierarchy theory, we classify the ordinary quantifier prefixes and sentences of higher-order logic as follows. The *character* of any universal quantifier is $\forall$, the *character* of the existential quantifier is $\exists$. A sequence of quantifiers is a Σ_n^m prefix if the highest order of any of its variables is $m+1$, its first quantifier is existential, and it has n quantifiers including the first which are of different character than their immediate predecessors. A Π_n^m prefix is the dual of a Σ_n^m prefix. A sentence is a Σ_n^m sentence (Π_n^m sentence) if it is of the form $Q\varphi$ where Q is a Σ_n^m prefix (Π_n^m prefix) and φ is a formula all of whose quantified variables have order $\leq m$."

We define Σ_1^1 as an extension of the language of first order logic as follows:

Definition 2.6.1 (Σ_1^1) *Let $\sigma = (\mathtt{C}, \mathtt{P}, \mathtt{F})$ be a first order signature. (See definition 2.1.1.) We add two new sets of* second order *variables:*

- *Function variables: for each $n \in \mathbb{N}$, infinitely many n-ary function symbols $f_1^{(n)}, f_2^{(n)}, \ldots$ (with for each i, n: $f_i^{(n)} \notin \mathtt{F}$).*
- *Predicate variables: for each $n \in \mathbb{N}$, infinitely many n-ary predicate symbols $X_1^{(n)}, X_2^{(n)}, \ldots$ (with for all i, n: $X_i^{(n)} \notin \mathtt{P}$).*

Let $\sigma' = (\mathtt{C}, \mathtt{P} \cup \{X_i^{(n)} | i \in \mathbb{N}, n \in \mathbb{N}\}, \mathtt{F} \cup \{f_i^{(n)} | i \in \mathbb{N}, n \in \mathbb{N}\})$ be the extension of σ with the new function and predicate variables. As usual (cf. the definition of $\mathcal{L}_{FOL}$ in section 2.1), the superscripts (n) are omitted whenever the arities of the variables are clear from the context.

Then $\Sigma_1^1(\sigma)$ is the collection of second order sentences of the form

$$\exists f_{i_1} \ldots \exists f_{i_k} \exists X_{j_1} \ldots \exists X_{j_m} \varphi,$$

where φ is a sentence of $\mathcal{L}_{FOL}^{\sigma'}$, containing no function symbols that are not in $\mathtt{F} \cup \{f_{i_1} \ldots f_{i_k}\}$, and no relation symbols not in $\mathtt{P} \cup \{X_{j_1} \ldots X_{j_m}\}$. (In other words: all first order variables are bound by first order quantifications in φ, and all second order variables in φ are bound by an 'initial' sequence of existential *second order quantifications.)*

As with $\mathcal{L}_{FOL}$, we will usually omit explicit reference to the first order signatures σ. When we say Φ is a Σ_1^1-sentence, we mean that $\Phi \in \Sigma_1^1(\sigma)$ for some first order signature σ. Again, signatures play a role only implicitly in the phrase 'suitable model': If Φ is a Σ_1^1-sentence, a first order model of signature σ is called suitable for Φ if $\Phi \in \Sigma_1^1(\sigma)$ (i.e. the model has an interpretation for all symbols in Φ).

For second order logic, and Σ_1^1 as part of it, the meta-logical properties depend largely on which domains of quantification are chosen for the second order variables: see [Vää01]. There are basically two approaches:

- *full* semantics, in which the set $A^{(A^n)}$ of n-ary functions on the domain A of a model $\mathfrak{A}$ (as determined by the axioms of ZFC) serves as domain of quantification for the n-ary function variables, and similarly, the set $\mathcal{P}(A^n)$ of all n-ary predicates on A serves as domain of quantification for all n-ary predicate variables. We use the turnstile '$\models_{\mathbf{sol}}$' for full second order semantics.

- Henkin-semantics, in which the domains of quantification for the second-order variables are given as explicit parameters in the models.

The first option, full semantics, allows us to evaluate second order sentences in first order models: the domains of quantification for the second order variables are determined in terms of the (individual) domain A, by the axioms of ZFC. The second option implies that in order to interpret second order formulas, first order models need to be extended with separate domains of quantification for the second order variables.

We refer to e.g. [Man96] and [Lei94] for more extensive descriptions of second order logic in general.

In the next chapter we will show how Hintikka uses Σ_1^1 to reason about strategies (and thereby: to reason about IF-logic).

Chapter 3

Skolemization and falsity conditions

We follow [Hin96] in regarding strategies in semantic games as sets of Skolem functions. First, we discuss and give a formal account of the Skolemization procedure which is central in Hintikka's approach to IF-logic, as it delivers Σ_1^1-*truth conditions* for IF-sentences. We notice that for soundness of the procedure, the models of evaluations need to satisfy some mild conditions; most notably, they should contain at least two elements.

Using a generalization of the IF-language, allowing us to write the negation of an IF-sentence into negation normal form, we give a procedure to obtain *falsity conditions* for IF-sentences. The expressive power of an IF-sentence is then shown to be captured in a stronger sense by a *pair* of Σ_1^1-sentences. We translate a recent result of John Burgess for Henkin sentences to show that, conversely, any pair of incompatible Σ_1^1-sentences corresponds with an IF-sentence.

The study of *falsity* gives rise to some reflections on the nature of game-theoretic negation. Furthermore, we explain how the order of the Skolemization steps (inside-out versus outside-in) makes a difference in IF-logic.

This chapter is based on [Dec05].

3.1 Introduction: what is a strategy?

In chapter 2, we gave a definition of the semantic games in terms of their rules. For the evaluation of (IF-)sentences however, it is the notion of *strategy* that is most crucial, and this was not formalized as a mathematical object in this definition.

In general understanding, a strategy prescribes a choice for a player in every possible position of the game in which it is that player's turn. In semantic games, we may describe every choice associated with a subformula by a function working on a set of valuations, and a strategy can be regarded as a sequence of such functions. The arguments of those functions reflect the available input information of the choices. By this correspondence, the existence of a winning strategy can be expressed as an existential second order sentence: a statement about the existence of functions satisfying certain first order properties. This *truth condition* can be obtained by the syntactical procedure of Skolemization.

In Hintikka's work, the connection between the functions constituting strategies in semantic games, and (generalized) Skolem functions is taken to be an immediate one (e.g. [Hin96, p. 40]). It is not made expplicit how this connection follows from the definition of semantic games in terms of their rules; we will give such a formalization in the next chapter (section 4.4). In [Hin96, p. 49], an argument is made in the converse direction: any existential second order statement can be interpreted as the winning condition for Eloïse in some well-interpreted semantic game. But this game may not be a semantic game for a *traditional* first order formula. It is enough to extend it to IF-sentences to let every existential second order sentence be a winning condition for a semantic game: an *inverse* Skolemization procedure translates any Σ_1^1-sentence into an IF-sentences of which it is a truth condition.

This approach is closely related to the theory of first order logic with Henkin quantifiers, introduced by Henkin in [Hen61]. Sentences in this language are also interpreted by Skolemization into Σ_1^1-sentences (with supporting motivation in terms of games), and the expressive power was proved to be equal to Σ_1^1 independently by Enderton in [End70] and by Walkoe in [Wal70].

In this chapter, we first give a precise account of the generalized Skolemization procedure to obtain Σ_1^1-truth conditions for IF-sentences, and the translation of Σ_1^1-sentences into IF-sentences that are true in the same models. In the book [Hin96], as in most other work on IF-logic, the focus is almost exclusively on the expressive power by truth conditions. In a second part, we show how *falsity* conditions can be formulated as Σ_1^1-sentences as well by some syntactical manipulations and the same Skolemization procedure. We end the chapter with some reflections on the nature of game-theoretic negation, and on the semantic assumptions underlying the (syntactic) translation procedures.

3.2 The generalized Skolemization procedure

3.2.1 Skolemization for first order logic

For first order logic with traditional semantics, Skolemization is often used as an instrument to eliminate existential quantifiers (e.g. in automated theorem proving, cf. [Fit96, section 7.11]). It replaces existentially quantified variables by 'functional' terms, explicitly expressing how these variables depend on the universal quantifications that have scope over it.

For simplicity (and as usual), we formulate Skolemization here for first order formulas that are in negation normal form. (Theorem 7.11.2 of [Fit96] is formulated for arbitrary first order formulas, using the concept of positive and negative occurrences of subformulas.)

Theorem 3.2.1 (Skolemization (steps) for first order logic) *Suppose φ is a first order formula in negation normal form, and let $\exists x\psi(x)$ be a subformula of φ. We indicate this by writing $\varphi[\exists x\psi(x)]$. Suppose $Fv(\psi(x))-\{x\} = \{y_1,\ldots,y_n\}$, and let f be an n-ary function symbol that does not occur in φ. Then the formula $\varphi[\exists x\psi(x)]$ is satisfiable if and only if the formula $\varphi[\psi(f(y_1,\ldots,y_n))]$ is.*

Proof: See the proof of Theorem 7.11.2 in [Fit96, p. 188]. $\triangleleft$

We call the process of removing one existential quantification and replacing the variable it bound by a functional term, as in the Theorem, a *Skolemization step*. We will use the term *Skolemization* for the removal of all existential quantifications in this manner.

Note that Skolemization is a non-deterministic procedure, even if we forget about the indeterminacy in the choice of the 'new' function symbols. Another type of indeterminacy results from the fact that the order of the removal of the existential quantifications is not prescribed. For example, the formula $\forall x\exists y\exists zR(x,y,z)$ becomes $\forall x\exists zR(x,f(x),z)$ after replacing the outer existential quantification $\exists y$, and if we then replace the inner quantification $\exists z$, we get the Skolemized form: $\forall xR(x,f(x),g(x))$. Applying the Skolemization steps 'inside-out', we first get $\forall x\exists yR(x,y,f(x,y))$, followed by $\forall xR(x,g(x),f(x,g(x)))$. The resulting formulas are visibly different: the inside-out procedure shows how the replaced variables depended on other existentially quantified variables, while the outside-in procedure only shows direct dependence on universally quantified variables. But the theorem above ensures that in any chosen order of Skolemization steps, the result of Skolemization for a first order formula is *equisatisfiable* with the original formula. We will show in section 3.9 how a similar result for IF-sentences fails.

In the form of Theorem 3.2.1, Skolemization is a statement on *satisfiability*: there is a (suitable) model that satisfies the original formula if and only if there is a model (with an extended signature) that satisfies the Skolemized formula. Skolemization steps reduce the number of quantifiers, which improves the efficiency of algorithms checking the logical validity of first order formulas: a first order formula ψ is valid if and only if $\neg\psi$ is not satisfiable, which is the case if and only if a Skolemization of $\neg\psi$ is not satisfiable. To check the latter algorithmically, we could apply Tableaux methods (see e.g. [Fit96, p. 189]).

To make Skolemization a matter of *truth* in a given model, we leave the first order level. If we treat the new function symbols as function *variables*, we can make the Skolemized formula into a second order formula of the same signature (without any free second order variables) as follows: we prefix the Skolemized formula with a second order quantification $\exists f_i$ for each newly introduced function f_i. The resulting existential second order sentence can then be read as a truth condition: a statement (in a different language) expressing the necessary and sufficient condition a model has to satisfy in order for the sentence to be true in that model. In the context of GTS, this Σ^1_1-Skolemization of a first order sentence φ may be read as expressing the existence of a winning strategy for Eloïse.

Theorem 3.2.2 (Skolemization into Σ_1^1) *Let $\varphi[\psi_1(f_{i_1}(\bar{y}_1)), \ldots, \psi_k(f_{i_k}(\bar{y}_k))]$ be the result of k successive applications of theorem 3.2.1 to the first order sentence φ, and if $\mathfrak{A}$ is suitable model for φ, then:*

$$\mathfrak{A} \models \varphi \textit{ iff } \mathfrak{A} \models_{\mathtt{sol}} \exists f_{i_1} \ldots \exists f_{i_k}[\varphi[f_{i_1}(\bar{y}_1), \ldots, f_{i_k}(\bar{y}_k)]]$$

Proof: This follows from Theorem 3.2.1 and the Axiom of Choice. $\triangleleft$

Note that if φ is a first order formula of signature σ, then the Σ_1^1-sentence Φ in the theorem:

$$\Phi := \exists f_{i_1} \ldots \exists f_{i_k}[\varphi[f_{i_1}(\bar{y}_1), \ldots, f_{i_k}(\bar{y}_k)]],$$

is also of signature σ. So, any model suitable for the first order sentence φ is suitable for the Σ_1^1-sentence Φ and conversely. Note that first order semantics ('$\models$') does not interpret the second order quantifications in the Σ_1^1-sentence Φ, therefore Φ has to be evaluated with second order semantics. With '$\models_{\mathtt{sol}}$', we indicate *full* semantics for second order logic.

This theorem depends on the use of full semantics, in which the axioms of ZFC determine the domain of quantification for the function variables (see section 2.6).

3.2.2 Skolemization for IF-sentences

Skolemization for IF-sentences (that are automatically in negation normal form by definition 2.5.1) is in two respects an extension of Skolemization for first order logic. First, the generalized procedure deals with the slash operator applied to existential quantifications (by omitting the variables occurring under the slash as arguments for the corresponding Skolem function). Second, it also introduces Skolem functions for the *disjunctions*. The latter is easily motivated by the fact that in the semantic games, disjunctions correspond to moves by Eloïse as well. More technically, Skolemization of the disjunctions is needed in order to remove slash operators applied to disjunctions in the IF-sentence, in order to get a truth condition that is in classical Σ_1^1, i.e. without any slashes.

We define the procedure of Skolemization for IF-sentences as follows (in Hintikka's book [Hin96] there is no formalized definition):

Definition 3.2.3 (Skolemization for IF-sentences) *Let φ be an IF-sentence. A* **Skolemization** *of φ is a Σ_1^1-sentence*

$$\exists f_{i_1} \ldots \exists f_{i_m} \varphi'[\psi_1(f_{i_1}(\bar{z}_1)), \ldots, \psi_m(f_{i_m}(\bar{z}_m))]$$

where the first order formula φ' is obtained by repetition of the following replacement steps until all existential quantifications and (original) disjunctions occurring in φ are replaced

(a) if $\exists y_{/x_1,\ldots,x_k}\psi(y)$ occurs as a subformula in φ under the scope of the universal quantifications in $\{\forall x_1, \ldots, \forall x_k, \forall z_1, \ldots, \forall z_n\}$, and if f is an n-ary function symbol that does not occur in φ, then replace $\exists y_{/x_1,\ldots,x_k}\psi(y)$ in φ by

$$\psi(f(z_1, \ldots, z_n))$$

(b) if $\psi_1 \vee_{/x_1,\ldots,x_k} \psi_2$ occurs as a subformula in φ under the scope of the universal quantifications in $\{\forall x_1, \ldots, \forall x_k, \forall z_1, \ldots, \forall z_n\}$, and if f is an n-ary function symbol that does not occur in φ, then replace $\psi_1 \vee_{/x_1,\ldots,x_k} \psi_2$ in φ by

$$(f(z_1, \ldots, z_n) = 0 \wedge \psi_1) \vee (f(z_1, \ldots, z_n) \neq 0 \wedge \psi_2)$$

(This procedure is also applied if $k = 0$, in other words: if the quantification or disjunction is not slashed.)

For example, the IF-sentence

$$\forall x \forall y \exists z_{/x}[R(x, y, z) \vee_{/y} P(x, y, z)]$$

is, by this Skolemization procedure, translated into the Σ_1^1-sentence

$$\exists f \exists g \forall x \forall y[(g(x) = 0 \wedge R(x, y, f(y))) \vee (g(x) \neq 0 \wedge P(x, y, f(y)))].$$

We make three remarks about this procedure. First, apart from the choice of function symbols, the procedure above is now deterministic: because existentially quantified variables are not taken as arguments for the Skolem functions, it does not matter anymore in which order the existential quantifiers are eliminated. (Not taking the existentially quantified variables as arguments corresponds to applying Theorem 3.2.1 outside-in.)

Second, note that this translation is strictly speaking not *signature preserving* (unlike the first order procedure 3.2.2): for the interpretation of the result of the second clause, a suitable model should interpret the constant '0' and the predicate symbol *equality* '='. Furthermore: not just any model with this extended signature will give the Skolemization the intended interpretation. In order to do so, a model needs to satisfy the axioms stating that the interpretation of '=' is in fact *equality* (and not an arbitrary binary predicate), and also, the model should contain at least two elements. We will call models that satisfy these requirements **IF-safe**:

Definition 3.2.4 (IF-safe models) *A first order model $\mathfrak{A}$ is called* **IF-safe** *if its signature σ contains the individual constant '0' and the binary relation symbol '=', and if it satisfies the equality axioms for $\mathcal{L}^\sigma_{FOL}$ (see [Fit96, section 8.3]) and $\exists x[x \neq 0]$ ($\mathfrak{A}$ contains at least two elements).*

Third, we realize that there is a hazard of circularity in clause *(b)*, as it replaces a disjunctive subformula (possibly with a slashed disjunction), by a new disjunctive formula (without a slashed disjunction). The procedure assumes that the new disjunction will not be replaced again. Formally, we could ensure this by only replacing slashed disjunctive subformulas ($\psi_1 \vee_{/x_1,\ldots,x_k} \psi_2$ with $k > 0$). However, this destroys the one-one correspondence between moves of Eloïse in the semantic game on the one hand, and Skolem-functions on the other.

As mentioned in the introduction to this section, Hintikka gives no (formal) argument for the following statement:

Statement 3.2.5 (Hintikka's approach) *If φ is an IF-sentence, and Φ is a Σ^1_1-sentence obtained by the Skolemization procedure of definition 3.2.3, then for any suitable* **IF-safe** *model $\mathfrak{A}$:*

$$\mathfrak{A} \models_{\mathtt{t}} \varphi \text{ iff } \mathfrak{A} \models_{\mathtt{sol}} \Phi$$

The requirement of 'IF-safety' is not explicit in the description in [Hin96]. But as we noted in our comments on definition 3.2.3, it is necessary in order for the translation procedure to be sound.

We have reasons not to state 3.2.5 as a Theorem, but refer to this as 'Hintikka's approach'. Hintikka takes the Skolemization procedure of definition 3.2.3, based on the 'outside-in' Skolemization procedure for first order logic, as *primitive* for the interpretation of IF-sentences. As he admits elsewhere ([Hin02a, p. 407], see the quote on page 50 of this thesis), the semantic games of GTS only serve as "an explication of our pretheoretical ideas about quantifiers and truth". This means that statement 3.2.5 is a definition rather than a theorem.

A consequence of taking this Skolemization procedure as primitive, is that dependencies of previous existential quantifiers are neglected. For first order logic, this happened to give the same result as taking these dependencies into account. In section 3.9 and in the next chapter we show that for some IF-sentences (like Hodges' example (1.7)), a natural game-theoretic interpretation of the semantic games would give a *different* truth condition from the one obtained by Skolemization as in definition 3.2.3. This interpretation would correspond to an 'inside-out' Skolemization procedure, taking dependencies of previous existential quantifications into account.

In response to Hintikka's attitude to semantic games as informal explications of our pre-theoretical ideas, we reply that we may need to form new pre-theoretical ideas for IF-logic. One of our pre-theoretical ideas about quantifiers, viz. that their dependence relation is transitive, is challenged by IF-logic. A principally game-theoretic interpretation can help us deal with this challenge (and we will try to do so in the next chapter).

3.2.3 Translating Σ^1_1-sentences to IF-sentences

In the previous subsection, we have shown how a Skolemization procedure translates IF-first order sentences into Σ^1_1-sentences, that are viewed as truth conditions for the IF-sentences. In this section we give a detailed formalization of Hintikka's procedure to translate a Σ^1_1-sentence into IF-first order logic, which is described in [Hin96, p. 62-63] (with an easily reparable mistake in (3.44): the double arrow should be a single one). This leads to the claim that IF-first order logic has precisely the expressive power of Σ^1_1.

First, it is proved that we may take any Σ^1_1-sentence Φ to be written in a specific form:

Lemma 3.2.6 *Any Σ^1_1-sentence Φ can be rewritten into a Σ^1_1-sentence Ψ in a specific form, such that for every* **IF-safe** *suitable model $\mathfrak{A}$:*

$$\mathfrak{A} \models_{\mathtt{sol}} \Phi \text{ iff } \mathfrak{A} \models_{\mathtt{sol}} \Psi,$$

This specific form is:

$$\exists f_1 \exists f_2 \ldots \exists f_l \forall x_1 \forall x_2 \ldots \forall x_n[\psi] \tag{3.1}$$

where ψ *is a* quantifier-free *first order formula (in which the function variables* $f_1, \ldots, f_l$ *and individual variables* $x_1, \ldots, x_n$ *occur), and such that the following conditions are satisfied:*

1. *the function variables* f_i *do not occur nested;*

2. *each function variable* f_i *occurs with only one sequence of arguments.*

Proof: Let Φ_0 be an arbitrary Σ_1^1-sentence, i.e. a first order formula φ_0 prefixed by a list of existential quantifications $\exists f_{j_1} \exists f_{j_2} \ldots$ over function variables, and possibly also a list of existential quantifications over predicate variables $\exists X \ldots$. We bring Φ_0 into the form of the lemma by a sequence of transformations.

First, replace the first order formula φ_0 with an equivalent first order formula φ_1 in prenex normal form (which implies negation normal form). Let φ_2 be the result of the Skolemization of φ_1 according to Theorem 3.2.1, such that all first order existential quantifications are removed. Let $f_{i_1}, \ldots, f_{i_k}$ be the Skolem functions introduced in the process. Let Φ be the formula we get from Φ_0 by replacing φ_0 with φ_2 (which is of the form $\forall x_1 \ldots \forall x_n[\psi]$, with ψ quantifier free), and prefixing the result with existential quantifications over the introduced Skolem functions: $\Phi := \exists f_{i_1}, \ldots, \exists f_{i_k} \exists f_{j_1} \exists f_{j_2} \ldots \exists X \ldots \varphi_2$.

To satisfy the two extra conditions of the lemma, we apply the necessary number of the following transformations to Φ:

- If Φ contains an existential quantification over an n-ary *predicate variable* X, it is replaced by a quantification over an n-ary function variable f, which acts as 'characteristic function' for the original predicate: every occurrence of the predicate X with arguments $(t_1, \ldots, t_n)$ in Φ, is replaced by the atomic formula $f(t_1, \ldots, t_n) \neq 0$. For example: $\exists X \forall x[X(x)]$ becomes $\exists f \forall x[f(x) \neq 0]$.

- If in Φ two function variables occur *nested*, i.e. in a term $f_i(t_1, \ldots, t_n)$ we have $t_k = f_j(t'_1, \ldots, t'_m)$ for some $k \in \{1, \ldots, n\}$, then the nested term t_k is replaced by a *new* individual variable x, and the first order part

$$\varphi(f_i(t_1, \ldots, t_k, \ldots, t_n))$$

of Φ is replaced by

$$\forall x[x \neq f_j(t'_1, \ldots, t'_m) \vee \varphi(f_i(t_1, \ldots, x, \ldots, t_n))]$$

(It would be more natural to write '$x = f_1(y) \rightarrow \varphi(f_2(x))$' rather than '$x \neq f_1(y) \vee \varphi(f_2(x))$'. But we will use the outcome of this procedure for translation into IF-logic where we have no satisfactory interpretation for implication in general. The formulation with negation and disjunction, translates directly into IF-logic, because the negation sign (implicit in $\neq$)

behaves like classical negation on atomic formulas.)
For example:

$$\exists f_1 \exists f_2 \forall y \varphi(f_2(f_1(y))),$$

becomes

$$\exists f_1 \exists f_2 \forall y \forall x[x \neq f_1(y) \vee \varphi(f_2(x))].$$

- If in Φ one function variable occurs *with different sequences of variables*, say $f_i(x_1, \ldots, x_n)$ and $f_i(y_1, \ldots, y_n)$ both occur in the first order part φ of Φ, we then add an existential quantification of a *new* n-ary function variable f_j, replace occurrences in φ of $f_i(y_1, \ldots, y_n)$ by $f_j(y_1, \ldots, y_n)$, and conjunct the result with the quantifier free formula

$$\neg(x_1 = y_1 \wedge \ldots x_n = y_n) \vee (f_i(x_1, \ldots, x_n) = f_j(y_1, \ldots, y_n))$$

For example:

$$\exists f_1 \forall x_1 \forall x_2 \varphi(f_1(x_1), f_1(x_2)),$$

becomes

$$\exists f_1 \exists f_2 \forall x_1 \forall x_2[(x_1 \neq x_2 \vee f_1(x_1) = f_2(x_2)) \wedge \varphi(f_1(x_1), f_2(x_2))].$$

(The same remark holds with respect to the avoidance of implication.)

The result of each these transformation is equivalent to the original (in IF-safe models). ◁

Note that the Σ^1_1-sentences that result from the Skolemization procedure 3.2.3 are of the form (3.1) and satisfy the two conditions of the lemma. In fact, all sentences of this form can be seen as the result of Skolemization for some IF-formula, as will be demonstrated by the translation procedure from Σ^1_1 into IF-logic.

Theorem 3.2.7 *Every Σ^1_1-sentence is equivalent (on the class of IF-safe models) to the Skolemization of some IF-sentence.*

Proof: Given an arbitrary Σ^1_1-sentence Φ, we transform it into the form (3.1) of the previous lemma (satisfying the two conditions):

$$\exists f_1 \exists f_2 \ldots \exists f_l \forall x_1 \forall x_2 \ldots \forall x_n[\psi(f_1, \ldots, f_l, x_1, \ldots, x_n)] \tag{3.2}$$

To transform the latter Σ^1_1-sentence into an IF-sentence, apply the following transformation to replace the quantification $\exists f_i$ by an IF-quantification $\exists y_{i/Z_i}$, for $i = 1, \ldots, l$: let $\bar{x}_i$ be the (unique) sequence of arguments of f_i, and let $Z_i := \{x_{i_1}, \ldots, x_{i_m}\}$ be the set of variables from $x_1, \ldots, x_n$ that do *not* occur in $\bar{x}_i$. Remove the quantification $\exists f_i$, and insert the IF-quantification $\exists y_{i/Z_i}$ directly left from the quantifier free part ψ of the formula (such that it falls under the scope of all universal quantifications). In ψ, replace every occurrence of the term

$f_i(\bar{x}_i)$ by the variable y_i. (Note that under Hintikka's approach, the order of these replacements does not matter.)

The result is an IF-sentence of the form:

$$\varphi := \forall x_1 \ldots \forall x_n \exists y_{1/Z_1} \ldots \exists y_{l/Z_l}[\psi(x_1, \ldots, x_n, y_1, \ldots, y_l)]$$

The Skolemization of φ following the procedure of 3.2.3 then gives us the formula (3.2) back (modulo choice of the second order variable names).

(Technical remark: note that the Σ^1_1-formula may contain implication signs, but in the formula (3.2) these only occur between quantifier free first order formulas. On that level, an implication $a \to b$ is equivalent to $\neg a \vee b$, also in IF-logic: at atomic level, game theoretical negation and classical negation coincide. We can therefore safely translate them.) ◁

Corollary 3.2.8 (Translation from Σ^1_1 to IF) *For every Σ^1_1-sentence Φ, there exists an IF-sentence φ such that in every suitable IF-safe model $\mathfrak{A}$:*

$$\mathfrak{A} \models_{\mathtt{sol}} \Phi \textit{ iff } \mathfrak{A} \mid\models_{\mathtt{t}} \varphi$$

Proof: By the previous theorem and statement 3.2.5. ◁

It follows that all Σ^1_1-definable properties can be expressed in IF-logic. To illustrate this, we show how a Σ^1_1-formula expressing *Dedekind infinity* can be translated into an IF-sentence:

Example 3.2.9: (Dedekind infinity expressed in IF-logic) Dedekind infinity defines infinity of a set A by the existence of an injective, non-surjective function $f : A \to A$. The Σ^1_1-sentence

$$\exists c \exists f \forall x \forall y[(f(x) = f(y) \to x = y) \wedge f(x) \neq c]]$$

is true precisely in models whose domain is Dedekind infinite. This formula is almost in the form of Ψ in Lemma 3.2.6, except that the function variable f occurs both with argument x and with argument y. If we also avoid the use of implication, we get the equivalent Σ^1_1-sentence:

$$\exists c \exists f \exists g \forall x \forall y[(x \neq y \vee f(x) = g(y)) \wedge ((f(x) \neq g(y) \vee x = y) \wedge f(x) \neq c)].$$

This Σ^1_1-sentence is the Skolemization of the following IF-sentence:

$$\forall x \forall y \exists z_{/x,y} \exists s_{/y} \exists t_{/x}[(x \neq y \vee s = t) \wedge ((s \neq t \vee x = y) \wedge s \neq z)].$$

◇

Note that by a back-and-forth translation from IF-logic to Σ^1_1, the IF-sentences return in a specific normal form: on the outside a block of (unslashed) universal quantifiers, then a block of existential quantifiers (with slashes), and in the scope of all these quantifiers a quantifier free, *slash free* formula. This is the strong Skolem normal form mentioned by Hintikka in [Hin96, p. 60, item (D)]. (In his supporting argument, Hintikka uses a prenex normal form for IF-sentences, without proof. In chapter 6 of this thesis, we give a detailed account of the conditions under which IF-sentences can be written into a *strongly* equivalent prenex normal form.) Note that in the translation procedures the slashed disjunctions disappear, which proves that they do not essentially contribute to the expressive power of IF-logic.

This concludes this section on the standard translations from IF-logic to Σ_1^1 and back, which form the basis of most of the metaproperties stated by Hintikka in [Hin96], items (A)–(G) on pages 59–63: the fact that its expressive power equals that of Σ_1^1, compactness, (downward) Löwenheim-Skolem (by which fact it cannot distinguish between countable and uncountable), separation, Beth definability (which follows from separation). Also, it is crucial in the definability of a truth predicate within the IF-language (chapter 6 of the book). In all cases only the truth conditions of IF-sentences play a role. The rest of this chapter is dedicated to the study of *falsity* conditions, and their relation to truth conditions.

3.3 Focus on truth

It is a well known fact that finite depth, two-player win-loss games of perfect information are determined, in the sense that one of the players has a winning strategy. This result follows from Zermelo's theorem in [Zer13], or the more general theorem of Gale and Stewart in [GS53] (and uses the axiom of choice). By this fact, it follows that GTS makes first order sentences always either true or false. For IF-sentences in general, this is no longer the case: the interpretation of the slash operator turns the semantic games into games of imperfect information (cf. section 2.5). In those games, it can be the case that none of the two players has a winning strategy. Hence: if an IF-sentence φ is not true in a given suitable model $\mathfrak{A}$, this does not necessarily imply that φ is false in that model.

In his book, as in his other work on the subject, Hintikka focuses strongly on when IF-sentences are true. We give three points where this focus is visible (these have also been pointed out by others in the literature, e.g. []). We illustrate these points by a comparison of the simple IF-formula $\forall x \exists y_{/x}[x = y]$ with the first order formula $\exists y \forall x[x = y]$.

The **first** point (also pointed out in [vB00a] and [vB05]) is the notion of **equivalence**: for Hintikka, sentences are equivalent if they are *true* in exactly the same models (cf. [Hin96, p. 65]):

Definition 3.3.1 (truth-equivalence) *For all IF-sentences φ and ψ:*

$$\varphi \equiv_t \psi \overset{d}{\Leftrightarrow} \text{for all suitable models } \mathfrak{A} : \mathfrak{A} \models_t \varphi \text{ iff } \mathfrak{A} \models_t \psi.$$

We call '$\equiv_t$' **truth equivalence**.

In this sense, the formulas $\forall x \exists y_{/x}[x = y]$ and $\exists y \forall x[x = y]$ are equivalent: they are both true in one-element models only. However, the first formula is never false (Abélard can never pick an element that is unequal to all elements), whereas the second formula is false in all models with at least two elements (Abélard, knowing the value chosen by Eloïse for y, can then pick a different value). This means that IF-sentences that have the same truth conditions, may have different falsity conditions. This shows that falsity conditions cannot be expressed in terms of truth conditions, and are hence a *separate* aspect of the descriptive power of an IF-sentence.

We prefer a stronger notion of equivalence, that also takes falsity into account. Such notion is a natural extension of truth-equivalence:

Definition 3.3.2 (strong equivalence) *For all IF-sentences φ and ψ:*

$$\varphi \equiv \psi \overset{d}{\Leftrightarrow} \varphi \equiv_{\mathfrak{t}} \psi \text{ and } \varphi \equiv_{\mathfrak{f}} \psi,$$

where $\equiv_{\mathfrak{f}}$ is the obvious counterpart of $\equiv_{\mathfrak{t}}$:

$$\varphi \equiv_{\mathfrak{f}} \psi \overset{d}{\Leftrightarrow} \text{ for all suitable models } \mathfrak{A} : \mathfrak{A} \models_{\mathfrak{f}} \varphi \text{ iff } \mathfrak{A} \models_{\mathfrak{f}} \psi.$$

We call '$\equiv_{\mathfrak{f}}$' **falsity equivalence**, *and '$\equiv$'* **strong equivalence**.

A **second** issue that makes the focus on truth evident, is related to the issue of equivalence. In the previous section, we have seen how every IF-sentence has a Σ_1^1-*translation*, and conversely, how every Σ_1^1-sentence is *translatable* into an IF-sentence. The word 'translation' implies some form of equivalence.

But the translation procedures are only **truth-preserving**, and disregard the falsity-aspects of the IF-sentences involved. This is demonstrated by the application of the back-and-forth translation to the simple first order formula $\exists y \forall x[x = y]$:

	IF		Σ_1^1
(1)	$\exists y \forall x[x = y]$	$\searrow$	
(2)			$\exists f \forall x[x = f]$
(3)	$\forall x \exists y_{/x}[x = y]$	$\swarrow$	

The sentences (1) and (3) share (2) as their truth condition, but where the original formula is false in all models containing at least two elements, the formula that arises after the back-and-forth translation is never false. This example shows in particular that the back-and-forth translation is not closed on the first order fragment of IF-logic, exactly by the fact that falsity conditions are not preserved.

Finally, the focus on truth is also visible in the **syntax** (see definition 2.5.1): Hintikka defines IF-logic from first order formulas in negation normal form, and then applies the slash operator only to existential quantifiers and disjunctions. This means that the information restrictions only restrict Eloïse in her choice of strategies, and therefore affect only the truth aspect of a formula.

The formula $\exists y \forall x_{/y}[x = y]$ –which is *strongly* equivalent to our IF-example $\forall x \exists y_{/x}[x = y]$: both are true in one element models only, and never false– is for example not an IF-formula in Hintikka's definition.

[Hin96] stresses the importance of the descriptive function of logic: characterization of classes of models by logical formulas. Usually this is done by investigating the class of models in which a formula is true (by the given semantics). In the case of IF-logic, as the following sections will demonstrate, the class of models in which an IF-sentence is false can not be expressed in terms of the class in which the sentence is true. This makes falsity a second dimension of the descriptive power of IF-logic, which is neglected if we, like Hintikka, only focus on truth.

In the rest of this chapter, we will study how to characterize the class of models in which an IF-formula is false. We have to mention that in the definitions of Sandu's Appendix to [Hin96], falsity *is* taken into account. In this Appendix, Sandu defines the IF-language through the definition of a class of standard (quantifier) prefixes and their duals, thereby generalizing the IF-language as defined by Hintikka. We will give our own, more direct, generalization.

3.4 Technicalities: symmetric syntax ($\mathcal{L}_{IFS}$)

In this section, we will define an extension of IF-logic ('IFS'), which is closed under (the game-theoretic) negation, i.e. if φ is an IFS-sentence, then so is $\neg\varphi$. In the way Hintikka defines the IF-sentences (definition 2.5.1), this is not the case: IF-sentences are defined from first order formulas already in negation normal form, and with the slash applied to existential quantifications and disjunctions only. This means that the moves of Abélard are always of perfect information. As a consequence, there is no IF-sentence that is strongly equivalent to the negated IF-sentence $\neg\forall x\exists y_{/x}[x = y]$ (which is never true, and false in one-element models only), as this would be the sentence $\exists x\forall y_{/x}[x \neq y]$, containing a slashed universal quantifier. On the other hand, game-theoretical semantics gives us a precise understanding of what this latter formula means.

The extended language of IFS-sentences that we define in this section, is designed to remove the asymmetry between the two players, induced by the choice of syntax in definition 2.5.1. We allow quantifiers and connectives that correspond to moves for Abélard, to be slashed for variables whose value was chosen by Eloïse. We design it such that negation normal form is no longer included in the definition, and yet with the property that any IFS-sentence is strongly equivalent to an IFS-sentence that is in negation normal form. (We should remark that the formalism of Sandu's Appendix to [Hin96], does allow for the symmetric treatment of the players and is closed under negation, but employs quantifier-blocks closely related to Henkin quantifiers, cf. [Hin96, p. 255]. We choose a more elementary language.)

Because we allow negation to occur anywhere in the formula, we need the following definition to keep track of which quantifiers and connectives correspond to moves of which player:

Definition 3.4.1 (polarity of quantifiers and connectives) *If φ is a first order sentence (not necessarily in negation normal form), and $Qx\psi$ is a subformula of φ, then Qx is called* **positive in** φ *if $Q = \exists$ and $Qx\psi$ is in the scope of an* *even number of negation signs in φ, or $Q = \forall$ and $Qx\psi$ is in the scope of an odd number of negation signs. In the two dual combinations ($Q = \forall$ and even, $Q = \exists$ and odd), Qx is called* **negative in** φ. *Similarly, if $\psi_1 \diamond \psi_2$ is a subformula of φ, then $\diamond$ is called* **positive in** φ *if $\diamond = \vee$ and $\psi_1 \diamond \psi_2$ is in the scope of an even number of negation signs in φ, or $\diamond = \wedge$ and $\psi_1 \diamond \psi_2$ is in the scope of an odd number of negation signs. In the two dual combinations ($\diamond = \wedge$ and even, $\diamond = \vee$ and odd), $\diamond$ is called* **negative in** φ. *Schematically:*

in φ	*even number* $\neg$*'s*	*odd number* $\neg$*'s*
$\exists, \vee$	*positive*	*negative*
$\forall, \wedge$	*negative*	*positive*

In sentences in negation normal form, the existential quantifications are precisely the positive quantifications, and the universal quantifications the negative ones. Similarly, the disjunctions are precisely the positive connectives, and the conjunctions the negative ones.

Definition 3.4.2 (symmetric IF-sentences: $\mathcal{L}_{IFS}$) *Given a first order signature σ, the set of IFS-sentences of signature σ: $\mathcal{L}^{\sigma}_{IFS}$, is determined by the following conditions:*

1. Every *first order sentence of signature σ (not necessarily in negation normal form) is in $\mathcal{L}^{\sigma}_{IFS}$.*

2. *Let φ be in $\mathcal{L}^{\sigma}_{IFS}$. Then the result of any of the following replacements is also in $\mathcal{L}^{\sigma}_{IFS}$:*

   (a) replacement of a positive quantification Qy in φ, occurring within the scope of negative quantifiers among which $Q_1 x_{1/Z_1}, \ldots, Q_k x_{k/Z_k}$, by $Qy_{/\{x_1,x_2,\ldots,x_k\}}$

   (b) replacement of a negative quantification Qy in φ, occurring within the scope of positive quantifiers among which $Q_1 x_{1/Z_1}, \ldots, Q_k x_{k/Z_k}$, by $Qy_{/\{x_1,x_2,\ldots,x_k\}}$

   (c) replacement of a positive connective $\diamond$ in φ, occurring within the scope of negative quantifiers among which $Q_1 x_{1/Z_1}, \ldots, Q_k x_{k/Z_k}$, by $\diamond_{/\{x_1,x_2,\ldots,x_k\}}$

   (d) replacement of a negative connective $\diamond$ in φ, occurring within the scope of positive quantifiers among which $Q_1 x_{1/Z_1}, \ldots, Q_k x_{k/Z_k}$, by $\diamond_{/\{x_1,x_2,\ldots,x_k\}}$

   (The notation $Q_i x_{i/Z_i}$ denotes an ordinary first order quantification, i.e. without a slash, if Z_i is taken to be the empty set.)

3. *This defines all $\mathcal{L}^{\sigma}_{IFS}$-sentences*

We call φ an IFS-sentence, if there is a first order signature σ such that $\varphi \in \mathcal{L}^{\sigma}_{IFS}$.

This language now includes the sentence $\exists y \forall x_{/y}[x = y]$, which is strongly equivalent to the IF-sentence $\forall x \exists y_{/x}[x = y]$.

Note that we did not extend the language of the IF-sentences with elements that require new semantics. Instead, we enable a more complete use of game theoretical semantics, in the definition of which no difference is made between the capacities of both players (we *emancipated* the syntax, to borrow a term we found in the 2002 version of [vB00b]). Also, because this language does not *presuppose* negation normal form (like the IF-language by definition 2.5.1), the negation rule gets truly used as role exchange. The definition of this symmetric language took some extra work because we adopted Hintikka's convention to make choices of one player in the game only independent of moves of his/her opponent (and not of his/her own, at least: not explicitly, cf. section 3.9). The syntactically determined polarity of the quantifiers and connectives, corresponds with the player who makes the associated moves in the semantic games for the formula: positive quantifiers and connectives give rise to moves for Eloïse, negative quantifiers and connectives give rise to moves for Abélard.

This enables us to state the following:

Lemma 3.4.3 (De Morgan's laws for $\mathcal{L}_{IFS}$) *For any IFS-sentence φ, write $\varphi[\psi] \equiv \varphi[\psi']$ to express that if ψ occurs as a subformula of φ, then the result of replacing ψ with ψ' in φ, is strongly equivalent to φ. We can then state De Morgan's laws for IF-logic as follows:*

$$\varphi[\neg(\forall x_{/Y}\psi)] \equiv \varphi[\exists x_{/Y}(\neg\psi)] \tag{3.3}$$
$$\varphi[\neg(\exists x_{/Y}\psi)] \equiv \varphi[\forall x_{/Y}(\neg\psi)] \tag{3.4}$$
$$\varphi[\neg(\psi_1 \wedge_{/Y} \psi_2)] \equiv \varphi[(\neg\psi_1) \vee_{/Y} (\neg\psi_2)] \tag{3.5}$$
$$\varphi[\neg(\psi_1 \vee_{/Y} \psi_2)] \equiv \varphi[(\neg\psi_1) \wedge_{/Y} (\neg\psi_2)] \tag{3.6}$$

In fact, these equivalences hold in the strongest sense possible. Not only does *there exist* a winning strategy for the formula on the left side iff *there exists* a winning strategy for the one on the right side (for a given player, in a given model). The following stronger situation is the case: a winning strategy for the formula on the left side *is* a winning strategy for the right side and vice versa. This is because the games for the sentences on the left side, are exactly the same games as those for the sentences on the right side. What we mean by 'the same game' becomes more precise in section 4.4 of the next chapter, where we model semantic games in game-theoretical terms. Informally: the exchange of roles combined with the flipping of the quantifier (or connective) cancels out.

Corollary 3.4.4 (NNF) *For every IFS-formula φ there exists a strongly equivalent IFS-formula in negation normal form.*

Proof: This follows from the previous lemma and the fact that double negation cancels: $\varphi[\neg\neg\psi] \equiv \varphi[\psi]$ (exchanging roles twice in a row brings the game to the original situation). $\triangleleft$

We now reflect on the question whether the introduction of imperfect information for Abélard (which is made possible by our move from IF-sentences to IFS-sentences), might increase Eloïse's winning potential. The answer to this is 'no':

Lemma 3.4.5 *An IFS-sentence φ in negation normal form is truth-equivalent with the IF-formula φ' resulting from φ after removing the slashes at universal quantifiers and conjunctions.*

E.g.: $\forall x \exists y_{/x} \forall z_{/y} \psi \equiv_{\mathfrak{t}} \forall x \exists y_{/x} \forall z \psi$. (This lemma corresponds to item (i) of the Corollary on p. 258 in the Appendix of [Hin96].)

Proof: The imperfect information does not reduce the set of actions for Abélard: in the most restricted case for Abélard, he has to play 'constant choices', i.e. choices that do not respond to earlier moves of Eloïse. But even in that case, Eloïse will never know in advance which constants Abélard chooses to play. Her strategy therefore still has to be winning for all possible actions of Abélard. (This is an informal argument, by the absence of a formal representation of semantic games. When we give the representation in the game-theoretic extensive form in chapter 4, we see that information sets of the other players are no parameter in the definition of strategy.) ◁

This lemma makes every IFS-sentence truth equivalent to an IF-sentence, which enables us to use the Skolemization procedure for IF-sentences (definition 3.2.3) to find Σ_1^1-truth conditions for IFS-sentences. We needed the slashes at universal quantifiers and conjunctions in the process of writing an IFS-formula in a (strongly) equivalent negation normal form. But after this process, we can ignore them again when looking for a truth condition. This way, we avoid obtaining an existential second order truth condition with slashed universal quantifiers and connectives (which would not be a Σ_1^1-formula because of these slashes).

The dual of the lemma is of course also true, but we will not need it: an IFS-sentence φ in negation normal form is falsity-equivalent to the IF-formula φ'' resulting from φ after removing the slashes for existential quantifiers; e.g. $\forall x \exists y_{/x} \forall z_{/y} \psi \equiv_{\mathfrak{f}} \forall x \exists y \forall z_{/y} \psi$.

3.5 Winning conditions for both players

The following simple observation shows how truth- and falsity conditions are two sides of the same coin:

Lemma 3.5.1 *A falsity condition for an IF(S)-sentence φ is a truth condition of the IFS-sentence $\neg\phi$:*

$$\mathfrak{A} \models_{\mathfrak{f}} \varphi \textit{ iff } \mathfrak{A} \models_{\mathfrak{t}} \neg\varphi.$$

This follows easily from the interpretation of negation as role exchange in GTS.

Corollary 3.5.2 *We can express the truth and falsity conditions of an IF(S)-sentence φ by two Σ_1^1-sentences, $\Phi_{\mathfrak{t}}$ and $\Phi_{\mathfrak{f}}$.*

To obtain $\Phi_{\mathtt{t}}$, we write φ in negation normal form (if necessary, by Corollary 3.4.4), then drop occurring slashes for universal quantifiers (by Lemma 3.4.5), and finally apply the Skolemization procedure to the result (definition 3.2.3). To obtain $\Phi_{\mathtt{f}}$, we do the same for $\neg\varphi$.

As an example, we formulate a falsity condition for an IF-sentence that expresses Dedekind infinity:

Example 3.5.3: (falsity of an IF-sentence expressing infinity) The truth condition of the following IF-sentence (slightly different from the formula in Example 3.2.9), easily shows that the sentence expresses the existence of a injective, non-surjective function. We give its truth- and falsity conditions (in the Σ_1^1-sentences we use double implication to improve readability):

$$\begin{array}{rcl}\varphi & := & \exists z \forall x \forall y \exists s_{/y} \exists t_{/x}[(x \neq y \vee s = t) \wedge ((s \neq t \vee x = y) \wedge (s \neq z))] \\ \Phi_{\mathtt{t}} & : & \exists f_1 \exists f_2 \exists f_3 \forall x \forall y[(x = y \leftrightarrow f_2(x) = f_3(y)) \wedge (f_2(x) \neq f_1)] \\ \neg\varphi & \equiv & \forall z \exists x \exists y \forall s_{/y} \forall t_{/x}[(x = y \wedge s \neq t) \vee ((s = t \wedge x \neq y) \vee (s = z))] \\ \Phi_{\mathtt{f}} & : & \exists f_1 \exists f_2 \forall z \forall s \forall t[\neg(f_1(z) = f_2(z) \leftrightarrow s = t) \vee s = z]\end{array}$$

We can now conclude that φ is false in $\mathfrak{A}$ precisely if $\mathfrak{A}$ is a 1-element model, because only in that case: $\mathfrak{A} \models_{\mathtt{sol}} \Phi_{\mathtt{f}}$. ◇

We remind the reader that any IF-sentence that is true in precisely the class of infinite models, must be undetermined in some finite models. If an IF-sentence would be false in (precisely) all finite models, its falsity condition would be a Σ_1^1-sentence expressing finiteness. This is impossible because Σ_1^1 is compact. This shows how we should be careful with the notions like finiteness and infinity (the one is the contradictory negation of the other), in combination with the game-theoretic negation and the associated notions of truth and falsity. Because IF-logic does not have the *contradictory* negation implicit in the term *in*finity, expressibility of infinity does not imply expressibility of finiteness.

3.6 IF-sentences correspond with Σ_1^1-pairs

With the procedure described above, we have a mechanism to associate with every IF-sentence φ, a pair $(\Phi_{\mathtt{t}}, \Phi_{\mathtt{f}})$ of Σ_1^1-sentences, such that for every suitable IF-safe model $\mathfrak{A}$, extending Hintikka's approach 3.2.5:

$$\mathfrak{A} \Vdash_{\mathtt{t}} \varphi \iff \mathfrak{A} \models_{\mathtt{sol}} \Phi_{\mathtt{t}}$$

and

$$\mathfrak{A} \Vdash_{\mathtt{f}} \varphi \iff \mathfrak{A} \models_{\mathtt{sol}} \Phi_{\mathtt{f}}.$$

The pair $(\Phi_{\mathtt{t}}, \Phi_{\mathtt{f}})$ divides the class of all suitable IF-safe models for φ into three disjoint subclasses: the class $\mathcal{M}_{\mathtt{t}}^{\varphi}$ of models satisfying $\Phi_{\mathtt{t}}$, the class $\mathcal{M}_{\mathtt{f}}^{\varphi}$ of models satisfying $\Phi_{\mathtt{f}}$, and finally the class $\mathcal{M}_{\mathtt{u}}^{\varphi}$ of models satisfying neither. We remark that the class $\mathcal{M}_{\mathtt{u}}^{\varphi}$ is not Σ_1^1- but Π_1^1-definable (Π_1^1 is *universal* second order logic, and the Π_1^1-definable classes are precisely the complements of the Σ_1^1-definable classes of models). If φ is first order, then of course $\mathcal{M}_{\mathtt{u}}^{\varphi}$ is empty.

For a Σ_1^1-sentence Φ, let $\mathcal{M}_{\mathtt{sol}}^{\Phi}$ denote the class of suitable models $\mathfrak{A}$ in which Φ is satisfied ($\mathfrak{A} \models_{\mathtt{sol}} \Phi$). We can then say that, if φ is an IF-sentence and Φ a Σ_1^1-sentence, then Φ is **a truth condition for** φ if $\mathcal{M}_{\mathtt{t}}^{\varphi} = \mathcal{M}_{\mathtt{sol}}^{\Phi}$. Also, Φ is **a falsity condition for** φ if $\mathcal{M}_{\mathtt{f}}^{\varphi} = \mathcal{M}_{\mathtt{sol}}^{\Phi}$.

On a sentential (propositional) level, we can formulate a nice correspondence between the propositional connectives $\wedge, \vee, \neg$ operating on IF-sentences (φ,ψ), and operations on the pairs of Σ_1^1-sentences as their truth- and falsity conditions:

IF-sentence	$\Sigma_1^1 \times \Sigma_1^1$
φ	$(\Phi_{\mathtt{t}}, \Phi_{\mathtt{f}})$
ψ	$(\Psi_{\mathtt{t}}, \Psi_{\mathtt{f}})$
$\neg\varphi$	$(\Phi_{\mathtt{f}}, \Phi_{\mathtt{t}})$
$\varphi \wedge \psi$	$(\Phi_{\mathtt{t}} \wedge \Psi_{\mathtt{t}}, \Phi_{\mathtt{f}} \vee \Psi_{\mathtt{f}})$
$\varphi \vee \psi$	$(\Phi_{\mathtt{t}} \vee \Psi_{\mathtt{t}}, \Phi_{\mathtt{f}} \wedge \Psi_{\mathtt{f}})$
$\neg\varphi \vee \psi$	$(\Phi_{\mathtt{f}} \vee \Psi_{\mathtt{t}}, \Phi_{\mathtt{t}} \wedge \Psi_{\mathtt{f}})$

The last line of the table gives some insight in the problems we have defining implication for IF-logic. It shows that a definition of implication in terms of game-theoretic negation and the disjunction would be stronger than intended. Intuitively, one expects an implication $\varphi \Rightarrow_{\mathcal{G}} \psi$ in GTS to express "if Eloïse has a winning strategy in $G_{\mathfrak{A}}(\varphi)$, then she has one in $G_{\mathfrak{A}}(\psi)$". This is a (material) implication on the level of strategies, and would correspond to the truth condition: $\neg\Phi_{\mathtt{t}} \vee \Psi_{\mathtt{t}}$. This condition is weaker than the condition $\Phi_{\mathtt{f}} \vee \Psi_{\mathtt{t}}$, because $\mathcal{M}_{\mathtt{f}}^{\varphi} \subseteq (\mathcal{M}_{\mathtt{t}}^{\varphi})^c$, which for IF-sentences in general is a proper inclusion. More importantly, $\neg\Phi_{\mathtt{t}} \vee \Psi_{\mathtt{t}}$ is in general not Σ_1^1, while truth conditions of IF-sentences are always Σ_1^1.

As suggested by [Hin96, 161, item (k)], this problem could be solved by moving to so called Extended IF-logic (EIF), in which contradictory (weak) negation ($-$) is added as a truth functional operator, which may be added sentence-initially:

$$\mathfrak{A} \mid\models_{\mathtt{t}} -\varphi \text{ iff } \mathfrak{A} \not\mid\models_{\mathtt{t}} \varphi.$$

In words: an EIF-sentence $-\varphi$ is true in a model $\mathfrak{A}$, if and only if Eloïse fails to have a winning strategy in the game $G_{\mathfrak{A}}(\varphi)$. But this introduces a syntactic element into the sentences that cannot be defined in terms of moves in the semantic games: it operates at the level of strategies. This makes the evaluation of EIF-sentences a 'higher order' activity. On the other hand, Hintikka's approach to IF-logic already uses the (existential) second order truth conditions as primitive, rather than the games. From that perspective a move to a bigger fragment of second order logic is not that big a step. Meta-logically, it would however mean a loss of some nice properties that IF-logic borrowed from Σ_1^1.

3.7 Σ_1^1-pairs corresponding with IF-sentences

In [CK99], the following question is posed:

> Is it true that for each pair of disjoint Σ^1_1-classes K_1, K_2 of structures of the same type there is an IF-sentence φ such that $K_1 = \mathcal{M}^{\varphi}_{\mathtt{t}}$ and $K_2 = \mathcal{M}^{\varphi}_{\mathtt{f}}$? In such case part (*b*) [of Theorem 5.1 in their paper] would give a quick proof of Craig's interpolation theorem. ([CK99, 30], translated into our notation)

In a recent note of J. Burgess, [Bur03], about this question in logic with Henkin quantifiers, the issue is attacked in the reverse direction. Using Interpolation for Σ^1_1, more specifically: the fact that for any incompatible pair of Σ^1_1-sentences Φ, Ψ (i.e. $\mathcal{M}^{\Phi}_{\mathtt{sol}} \cap \mathcal{M}^{\Psi}_{\mathtt{sol}} = \emptyset$), there is a first order sentence θ such that $\mathcal{M}^{\Phi}_{\mathtt{sol}} \subseteq \mathcal{M}^{\theta}_{\mathtt{sol}}$ and $\mathcal{M}^{\Psi}_{\mathtt{sol}} \subseteq (\mathcal{M}^{\theta}_{\mathtt{sol}})^c$, Burgess proves that for any pair of incompatible Henkin sentences (Φ_0, Φ_1), there is a Henkin sentence Θ such that Θ is (truth-)equivalent with Φ_0 and $\neg\Theta$ is (truth-)equivalent with Φ_1. In order to prove this, he needs a restriction similar to our restriction to the class of IF-safe models:

> "To avoid trivialities, in the logic of first-order sentences it is conventional to exclude models with an empty domain, while in the logic of Henkin sentences it will be convenient to exclude models with a one-element domain as well." ([Bur03, p. 2])

Note that under this restriction, the formula $\theta_0 := \forall x \exists y_{/x}[x = y]$ is *never* (i.e. in no IF-safe model) true, nor false. So: $\mathcal{M}^{\theta_0}_{\mathtt{t}} = \mathcal{M}^{\theta_0}_{\mathtt{f}} = \emptyset$.

We can easily translate the result for Henkin sentences into the following theorem about pairs of Σ^1_1-sentences and IF-logic:

Theorem 3.7.1 *Let Φ and Ψ be incompatible Σ^1_1-sentences. Then there is an IF-sentence θ such that Φ is a truth condition for θ, and Ψ is a falsity condition for θ.*

Proof: We use the IF-sentence $\theta_0 := \forall x \exists y_{/x}[x = y]$.

First, consider the simple case where $\mathcal{M}^{\Psi}_{\mathtt{sol}} = \emptyset$ (Ψ is an inconsistent Σ^1_1-sentence). Applying Corollary 3.2.8 to Φ, we find an IF-sentence φ such that $\mathcal{M}^{\varphi}_{\mathtt{t}} = \mathcal{M}^{\Phi}_{\mathtt{sol}}$. Now take for $\theta := \varphi \vee \theta_0$. This gives us:

$$\mathcal{M}^{\theta}_{\mathtt{t}} = \mathcal{M}^{\varphi \vee \theta_0}_{\mathtt{t}} = \mathcal{M}^{\varphi}_{\mathtt{t}} \cup \mathcal{M}^{\theta_0}_{\mathtt{t}} = \mathcal{M}^{\varphi}_{\mathtt{t}} \cup \emptyset = \mathcal{M}^{\varphi}_{\mathtt{t}} = \mathcal{M}^{\Phi}_{\mathtt{sol}}$$

$$\mathcal{M}^{\theta}_{\mathtt{f}} = \mathcal{M}^{\varphi \vee \theta_0}_{\mathtt{f}} = \mathcal{M}^{\varphi}_{\mathtt{f}} \cap \mathcal{M}^{\theta_0}_{\mathtt{f}} = \mathcal{M}^{\varphi}_{\mathtt{f}} \cap \emptyset = \emptyset = \mathcal{M}^{\Psi}_{\mathtt{sol}}$$

In the general case, we can find IF-sentences φ', ψ' such that $\mathcal{M}^{\varphi'}_{\mathtt{t}} = \mathcal{M}^{\Phi}_{\mathtt{sol}}$ and $\mathcal{M}^{\psi'}_{\mathtt{t}} = \mathcal{M}^{\Psi}_{\mathtt{sol}}$. We now form $\varphi := \varphi' \vee \theta_0$ and $\psi := \psi' \vee \theta_0$. Then: $\mathcal{M}^{\varphi}_{\mathtt{t}} = \mathcal{M}^{\Phi}_{\mathtt{sol}}$, $\mathcal{M}^{\psi}_{\mathtt{t}} = \mathcal{M}^{\Psi}_{\mathtt{sol}}$, *and* $\mathcal{M}^{\varphi}_{\mathtt{f}} = \mathcal{M}^{\psi}_{\mathtt{f}} = \emptyset$. By Craig's theorem for Σ^1_1, because Φ and Ψ are incompatible, there is a first order sentence χ such that $\mathcal{M}^{\varphi}_{\mathtt{t}} = \mathcal{M}^{\Phi}_{\mathtt{sol}} \subseteq \mathcal{M}^{\chi}_{\mathtt{t}}$ and $\mathcal{M}^{\psi}_{\mathtt{t}} = \mathcal{M}^{\Psi}_{\mathtt{sol}} \subseteq (\mathcal{M}^{\chi}_{\mathtt{t}})^c = \mathcal{M}^{\chi}_{\mathtt{f}}$ (χ is first order). Now take $\theta := \varphi \wedge (\neg\psi \vee \chi)$, and we will have the following, to complete the proof:

$$\mathcal{M}^{\theta}_{\mathtt{t}} = \mathcal{M}^{\varphi}_{\mathtt{t}} \cap (\mathcal{M}^{\neg\psi}_{\mathtt{t}} \cup \mathcal{M}^{\chi}_{\mathtt{t}}) = \mathcal{M}^{\varphi}_{\mathtt{t}} \cap (\mathcal{M}^{\psi}_{\mathtt{f}} \cup \mathcal{M}^{\chi}_{\mathtt{t}}) = \mathcal{M}^{\varphi}_{\mathtt{t}} \cap \mathcal{M}^{\chi}_{\mathtt{t}} = \mathcal{M}^{\Phi}_{\mathtt{sol}}$$

$$\mathcal{M}_{\mathtt{f}}^{\theta} = \mathcal{M}_{\mathtt{f}}^{\varphi} \cup (\mathcal{M}_{\mathtt{f}}^{\neg\psi} \cap \mathcal{M}_{\mathtt{f}}^{\chi}) = \mathcal{M}_{\mathtt{t}}^{\psi} \cap \mathcal{M}_{\mathtt{f}}^{\chi} = \mathcal{M}_{\mathtt{sol}}^{\Psi}$$

$\triangleleft$

We can conclude that the (strong) expressive power of IF-sentences consists of all pairs of incompatible Σ_1^1-sentences.

As Burgess stresses in [Bur03, 3], contrary negation on Henkin sentences (for which we can read game-theoretic negation on IF-sentences) "does not correspond to the semantic operation of complementation on classes of models, but further it does not correspond to any semantic operation at all." The theorem above shows that there can be many different IF-sentences that are truth equivalent, but pairwise not strongly equivalent. In fact, there are as many of these sentences as there are Σ_1^1-definable classes of models disjoint with the (one) class of models in which they are all true.

To illustrate this, for every $n \in \mathbb{N}$, let ψ_n be a first order formula expressing that the domain contains at least n elements (for example, we can take $\psi_3 = \exists x_1 \exists x_2 \exists x_3 [x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_3 \neq x_1]$), and let φ be an IF-formula that is true in infinite models only (e.g. the formula φ on page 38 of this thesis). Define $\varphi_n := \varphi \wedge \psi_n$. Then for every n, $\mathcal{M}_{\mathtt{t}}^{\varphi_n} = \mathcal{M}_{\mathtt{t}}^{\varphi}$ (i.e. the class of all infinite models), while for all m with $n < m$: $\mathcal{M}_{\mathtt{f}}^{\varphi_n} \subsetneq \mathcal{M}_{\mathtt{f}}^{\varphi_m}$ (i.e. the classes of models with less than n and m elements respectively).

3.8 Reflections on game-theoretic negation

In the last sections, we have focused on falsity aspects of IF-sentences, a topic closely related to the subject of game-theoretic negation (as witnessed by lemma 3.5.1). In this section we make some remarks on the nature of game-theoretic negation. In these remarks, we want to apply negation at non-atomic level, so we have to formulate the remarks for IFS-logic.

The definition of game-theoretic negation as role exchange, makes the formula $\varphi \vee \neg\varphi$ a statement expressing the determinacy of the games $G_{\mathfrak{A}}(\varphi)$. We know that games for IFS-sentences φ are in general not determined, so $\varphi \vee \neg\varphi$, the 'law' of the excluded middle, is not a logical law for IFS-logic.

This reminds us, of course, of negation in intuitionistic logic, which is also too strong to make the law of the excluded middle hold. We would like to point out that this eye-catching shared property does not automatically make IFS-logic and intuitionistic logic related. In this respect, we make three remarks.

First, as easily as we see a shared property of the negations in the two logics, we can also find a significant difference. In intuitionistic logic, the law of the excluded middle can be seen to be equivalent to the cancellation of double negation ($\neg\neg A \rightarrow A$), and conversely. In IFS-logic however, $\neg\neg\varphi$ is clearly equivalent to φ, by the fact that a double role exchange ("turning the game board around twice in a row") has no significant effect to the semantic game. Another inference scheme that distinguishes them, is the part '$\neg\forall x\psi \Rightarrow \exists x\neg\psi$' of De Morgan's laws, which does not hold intuitionistically.

Another important difference is that at the atomic level, game-theoretic negation is equal to classical negation: the winner of the game is determined by a purely classical evaluation of the atomic formula in the given model. This gives every run of a game a winner. In intuitionistic logic, it is possible that we do not have a proof for an atomic proposition, nor a proof that there can be no proof: atomic formulas may be undetermined.

Furthermore, game-theoretic negation can be seen to coincide with classical negation not only at the atomic level, but on all unslashed formulas: for all classical first order sentences φ:

$$\mathfrak{A} \models \varphi \overset{AC}{\rightleftarrows} \mathfrak{A} \mid\models_t \varphi, \text{ and:}$$

$$\mathfrak{A} \models \neg\varphi \overset{AC}{\rightleftarrows} \mathfrak{A} \mid\models_t \neg\varphi$$

This follows from the axiom of choice, because this makes first order sentences (interpreted classically) equivalent to their Skolemizations (as in Theorem 3.2.2), which are the truth conditions of these sentences in GTS. In the game-theoretic definition of truth, the axiom of choice is incorporated ([Hin96, 40]).

We see game-theoretic negation and its behavior in GTS for IFS-logic, as an extension of classical negation rather than related to intuitionistic negation. The truth value gap introduced by the extension of the game-theoretic negation beyond first order logic is in some sense unavoidable: truth and falsity conditions are incompatible and both Σ^1_1, while Σ^1_1 is not closed under classical negation (in other words: the collection of Σ^1_1-definable classes of models is not closed under complements). However, if we would like to regard game-theoretic negation as an extension of the *operation* that classical negation defines on the first order definable classes of models to the Σ^1_1-definable classes, we have to realize that the game-theoretic negation does not correspond to any operation at all, as shown by the result of Burgess (our theorem 3.7.1).

3.9 Skolem functions and strategies

In this chapter, we have followed Hintikka's approach, identifying strategies in semantic games with series of Skolem functions. These functions were obtained for IF-sentences by generalizing the Skolemization procedure for first order logic. In the original procedure, as we noted with respect to Theorem 3.2.1, the order of the Skolemization steps determined whether existentially quantified variables would occur as arguments of the Skolem functions. However, any chosen order gives an equisatisfiable result, and the existential second order closures are equivalent. The outside-in order, corresponding with universally quantified variables only as arguments, gives the least complex Skolemization, so this is how the procedure is usually applied.

For the interpretation of Henkin quantifiers (cf. section 1.4), the arguments of the Skolem functions are naturally taken to be universally quantified variables as well: the Henkin-sentence

$$\begin{pmatrix} \forall x & \exists y \\ \forall z & \exists u \end{pmatrix} R(x, y, z, u)$$

is interpreted through its Skolemization:

$$\exists f \exists g \forall x \forall z R(x, f(x), z, g(z)) \tag{3.7}$$

This is also Hintikka's intended truth condition for the (linear) IF-version, of the Henkin sentence above:

$$\forall x \forall z \exists y_{/z} \exists u_{/x} R(x, y, z, u). \tag{3.8}$$

But this linear notation makes one existential quantification ($\exists u$) fall in the scope of the other ($\exists y$), while the 2-dimensional notation of the Henkin quantifier makes clear that neither of the two existential quantifications is in the scope of the other. If we would base the Skolemization procedure for IF-sentences on 'inside-out' Skolemization for first order logic, we would get

$$\exists f \exists g \forall x \forall z R(x, f(x), z, g(z, f(x))) \tag{3.9}$$

as truth condition for the IF-sentence (3.8).

The linear IF-sentence (3.8) gives the order of the moves in the semantic game. When Eloïse chooses a value for u, values for x, z, *and* y have already been chosen. The slash operator forbids her to use the value x, so one would expect her to base her choice on the values of both z and y (while y was chosen as a function of the value of x: $f(x)$). This is reflected in the Skolemization (3.9), but not in (3.7), where the existentially quantified variable y (replaced by $f(x)$) does not occur as argument of g. We now wonder: are the two approaches equivalent?

That this is not the case, is demonstrated by the example proposed by Hodges ([Hod97a, p. 548]):

$$\forall x \exists z \exists y_{/x} [x = y] \tag{3.10}$$

Note that this formula is not expressible with Henkin quantifiers: the dependency of the quantifications in Henkin quantifiers forms a partial order, hence is **transitive**. In (3.10), we see that dependence of quantifications in IF-sentences may be **non-transitive**: y depends on z, z depends on x, but y may not depend on x.

Even though z does not occur in the atomic part of this IF-sentence, the move for Eloïse that corresponds to $\exists z$ in the semantic game is everything but inessential. Without the 'empty' quantification over z, in games for the IF-sentence $\forall x \exists y_{/x}[x = y]$, it is easy to see that Eloïse has no winning strategy in models containing more than one element. But with the additional move corresponding to the quantification $\exists z$, Eloïse gets the possibility to *signal* information that later is supposed to be hidden for her, which gives her a winning strategy in all models: she can choose the value for z equal to the value assigned to x, and then choose the value for y equal to the value assigned to z. This is reflected in the 'inside-out' truth condition for (3.10):

$$\exists f \exists g \forall x [x = f(g(x))]$$

and not in the truth-condition we get from the 'standard' Skolemization procedure of definition 3.2.3 (where f comes out as Skolem constant, i.e. a 0-ary Skolem function):

$$\exists f \exists g \forall x [x = f]$$

Hintikka intended to generalize the usual Skolemization procedure for first order logic (or stay as close as possible to the theory of Henkin quantifiers), as reflected by what we called 'Hintikka's approach' 3.2.5. He ensures this by introducing a convention:

> "At this point, it is in order to look back at the precise way the information sets of different moves are determined in semantic games with IF-sentences. The small extra specification that is needed, is that moves connected with existential quantifiers are always independent of earlier moves with existential quantifiers. [...] The reason for this provision is that otherwise 'forbidden' dependencies of existential quantifiers on universal quantifiers could be created through the mediation of intervening existential quantifiers." ([Hin96, p. 63])

(The quote does not mention explicitly that moves connected with *disjunction* are always independent of earlier moves with existential quantifiers, but we take that this is the case as well.) In [Jan02] this 'specification' is referred to as the **slashing convention**, and we will adopt this term for it.

In [Hod97a] and [CK99], we find alternative semantics for (more general) IF-languages, without this convention. The strategy functions in these semantics can be seen to correspond with inside-out Skolemization, as they use all previously quantified variables as arguments for the functions.

On sentences such as Hodges' formula (3.10), these semantics therefore differ from the evaluation under Hintikka's approach. We note that semantics corresponding to inside-out Skolemization, also give us the means to characterize Dedekind infinity in a more compact way (with less quantifiers than in examples 3.2.9 and 3.5.3): the IF-sentence

$$\exists u \forall x \exists y \exists z_{/x} [x = z \wedge y \neq u] \qquad \text{([CK99, example 1.4])}$$

gets, by inside-out Skolemization, the truth condition:

$$\exists k \exists g \exists f \forall x [x = f(k, g(k, x)) \wedge g(k, x) \neq k]$$

This Σ_1^1-sentence expresses that there is a function $g_k = g(k, \dots)$ that is injective (because it has a left-inverse $f_k = f(k, \dots)$), and non-surjective (because k is outside the range of g_k).

Dropping the slashing convention (or in terms of Skolemization, switching from outside-in to inside-out Skolemization) still delivers Σ_1^1-truth conditions. The expressive power of IF-logic as a whole therefore does not depend on the order chosen (see also [CK99, theorems 4.2 and 4.3]).

In order to make the slashing convention less counterintuitive, we may think of the players as being two **teams**, with each team consisting of as many players as there are positive (resp. negative) quantifiers and connectives in the formula. Each player then plays according to *one* function. They may communicate before the start of the game, to coordinate the choice of functions (whose arguments are exclusively choices of the opponent), but during the game, they no longer interact. This makes it easier to understand how for example the first order formula

$$\exists x \exists y[x = y] \tag{3.11}$$

can still be true in all (suitable) models, if Eloïse is supposed to have forgotten the value she chose for x when she has to choose the matching value for y. (This formula is one of the examples of [Jan02] to show how IF-logic sometimes conflicts with intuitions on *independence*.) In this *coordination game*, the two players of team Eloïse have to play constants, as there are no choices of the opponent in this game. But in a coordination phase before the start of the game, we can imagine them deciding to pick the same constant.

An interesting elaboration of this team-interpretation in game-theoretic terms, using the concept of *weak dominance*, is suggested in [Sev04].

Avoiding the slashing convention does not rid us of the problem of understanding game-theoretically how the formula (3.11) can be true. In the more general languages of [CK99] and [Hod97a], where any set of variables may occur under the slash (not only those whose values were chosen by the opponent), without the slashing convention, we can express a formula with the same 'meaning' by adding an explicit indendence to the second existential quantifier:

$$\exists x \exists y_{/x}[x = y]$$

This formula then has the same truth condition: $\exists f \exists g[f = g]$, hence is true in all models. A standard game-theoretic analysis using *information sets* to model the imperfect information, will also make this formula true in all models, as we will see in the next chapter. Our struggle to understand it, may be explained by the fact that the semantic games for these formulas are games of *imperfect recall*, which are conceptually difficult and are therefore even avoided in most work in Game Theory. We will come back to this in section 4.7 in the next chapter, where we formalize the semantic games in so-called *extensive form*.

In that chapter, we will start working with the more general language $\mathcal{L}_{IFG}$, which is the language of [CK99]. From then, we use the terms "IF-logic" or "IF-sentence" to imply that the slashing convention is adopted, while "IFG-logic" or "IFG-formula" imply that all independences are explicitly written in the formula(s).

3.10 Conclusions

We started this chapter by giving a detailed description of the Skolemization procedure for first order logic. By prefixing the result with existential quantifications over the newly introduced function symbols, the result of the Skolemization procedure is a Σ^1_1-sentence, that allows us to 'lift' Skolemization from being about *equisatisfiability* to *equivalence*.

We noted that the steps in the first order procedure can be applied in any order, giving different but equivalent results. If we apply the procedure 'outside-in', the Skolem functions only have universally quantified variables as argument, if we apply it 'inside-out', they also have existentially quantified variables as arguments. Hintikka generalizes the outside-in Skolemization procedure to IF-logic, and states that the result serves as *truth condition* for the original IF-sentence. We noted that the generalized Skolemization procedure for IF-logic presupposes a restriction to *IF-safe* models, i.e. models with equality and an interpretation for the constant '0', *and* containing at least two elements.

We also gave a precise presentation of Hintikka's procedure to translate any Σ^1_1-sentence back into an IF-sentence. The combination of the Skolemization procedure and this translation back, shows that the expressive power of IF-logic equals that of Σ^1_1. The result of the back and forth translation has a specific normal form (the Skolem normal form mentioned by Hintikka on [Hin96, p. 60]), which does not contain any slashed disjunctions. This shows that they do not essentially contribute to the expressive power. Furthermore, the back and forth translation is not closed on the first order fragment, which demonstrates that the procedure ignores the falsity aspects of the original formula.

In general, Hintikka focuses on truth of IF-sentences only, as demonstrated for example by his use of *truth equivalence* rather than *strong equivalence*. However, falsity of IF-sentences can be seen to be an independent second dimension to the descriptive power of an IF-sentence. We showed how a falsity condition for an IF-sentence can be obtained as a truth condition for its negation. In the process, we had to extend the original IF-language just enough to be able to write the negation of an IF-sentence into negation normal form. With both truth- and falsity conditions, IF-sentences are *strongly* characterized by a *pair* of Σ^1_1-sentences. Translation of a result of [Bur03] for Henkin quantifiers, shows that there is a correspondence between IF-sentences and incompatible pairs of Σ^1_1-sentences.

This demonstrates the curious fact that game-theoretic negation does not correspond with any operation on classes of models. Furthermore, we argued that game theoretical negation should not be associated with intuitionistic negation too easily, just because both make the 'law' of the excluded middle fail. IF-logic with GTS incorporates some nonconstructive elements (most notably $\neg\neg\varphi \equiv \varphi$) that make the game-theoretic negation more classical than intuitionistic.

The implication $\Phi \rightarrow \Psi$ for two Σ^1_1-sentences Φ, Ψ is not Σ^1_1 in general. Therefore, a natural notion of implication $\varphi \rightarrow \psi$ for IF-sentences φ, ψ (expressing that *if* Eloïse has a winning strategy in the game for φ, *then* she has one in the game for ψ) is not definable using the connectives $\vee, \wedge$ and the game-theoretic negation $\neg$. Adding a *sentence-initial* contradictory (weak) negation '$-$' ($-\varphi$ is true iff φ is not true) would solve this, but this introduces a connective that cannot be interpreted on the level of the game rules (only on the level of strategies).

Skolemization (outside-in) for first order logic, and the Skolemization procedure that constitutes the meaning for Henkin quantifiers, appear to have been Hintikka's inspirations for the semantics for IF-logic rather than the game theoretical semantics he advocates. However, in some cases, like for Hodges' formula (3.10), a straightforward game-theoretic analysis gives a different evaluation than the truth condition obtained by Hintikka's Skolemization procedure. This difference seems to be caused by the fact that in IF-logic, other than in first order logic or logic with partially ordered quantification, the dependence relation of the quantifiers can fail to be transitive. To remove this difference, Hintikka introduces a provision, which we call the *slashing convention*, that forces the game-theoretic interpretation to become the same as the interpretation by the Skolemization procedure.

In the next chapter, we approach IF-logic from the other side: we study game theoretical semantics from a game-theoretic perspective.

Chapter 4

Game theory as formal framework

From this chapter, we use the generalization $\mathcal{L}_{IFG}$ (from [CK99]) rather than Hintikka's language $\mathcal{L}_{IF}$, as the generalized language is more natural in the game-theoretic approach.

In the first part of this chapter, we present the terminology from Game Theory that is necessary to formalize semantic games as so-called games *in extensive form* (of imperfect information). We give a detailed account of how the components of this representation can be expressed in terms of the parameters of a semantic game (viz. an IFG-sentence and a suitable model), extending similar formalizations for IF-propositional logic in [SP01] and [SP03]. This gives a game-theoretically supported formalization of the concept of *strategy* in semantic games. We reflect on some characteristics of semantic games. Elaborating on an observation in [Hod97a], we demonstrate that, in general, independence of *connectives* instead of variables is game-theoretically problematic.

In the second, more exploratory part, we use the fact that we have 'embedded' semantic games into game theory: we attempt to apply game theoretical results and concepts to semantic games. This approach is closely related to ideas presented in van Benthem's program of exploring and exploiting the interplay between logic and game theory. We show what types of *imperfect recall* are present in semantic games, and we explore to which extent the four *Thompson transformations* on games in extensive form, correspond with logical equivalence schemes.

This chapter combines and extends the papers [Dec02] and [Dec04].

4.1 Introduction: using game theory in logic

In chapter 2, game theoretical semantics was defined by the description of semantic games in terms of their rules, and defining truth in terms of Eloïse having a winning strategy in such game. But Hintikka's formal arguments use the truth conditions resulting from the syntactic procedure of Skolemization, which we discussed in the previous chapter. The Skolemization procedure in itself bears no reference to the games and game theory brought in by Game Theoretical Semantics. Indeed, Hintikka seems to have intended the reference to games to be pre-theoretical rather than a formal basis for the logic:

> "Game theoretical semantics might at first look like a recondite way of approaching IF first-order logic. In reality, it is little more than an explication of our own pre-theoretical ideas about quantifiers and truth. For if you contemplate what Skolem functions are — and even if you do not contemplate it — obviously their job description is to produce the very "witness individuals" (usually depending on earlier choices of such witness individuals) that show the truth of the quantifier sentence in question. [...] One can therefore look upon game-theoretical semantics as being little more than an elaboration of our pre-theoretical idea of truth as applied to quantified sentences." [Hin02a, p. 407]

In his skeptical paper about the helpfulness of logical games in semantics [Hod01a], Hodges asks the question why these are generally considered "to 'shed light' ", or " 'help to understand' " (loc. cit., p. 18). We share the general skepticism of this paper about whether existing logics are *really* clarified by new semantics in terms of games. But in the case of IF-logic, we feel the *need* to understand the logic in terms of the semantic games of imperfect information that are defined to form its semantics. Especially because several examples (like Hodges' formula (3.10), or the other examples in the section Deathtraps of [Hod97a], show us that an analysis in terms of the games may differ from an analysis by Skolem-functions (while both are supposed to reflect our pre-theoretical ideas about quantifiers).

In this chapter we will deviate from the treatment of semantic games and strategies as pre-theoretical background, and study how (or even: if) we can reason about IF-logic if we take game theory as a formal framework. In order to do so, we model semantic games as games in extensive form.

So, our initial motivation for the modeling of semantic games in a suitable game theoretical framework, is to find a definition of the concept of strategy for semantic games as game theorists would give it in their standard framework, in order to bridge the conceptual gap between the description of semantic games in terms of their rules, and the concept of strategy as a set of Skolem functions.

While Hintikka seems to avoid using game theoretic concepts or results, the description of semantic games as extensive games has become more common in the work on IF-logic by other people, maybe most prominently in that of Johan van Benthem (early references being [vB00a] and [vB00b]). For IF-propositional logic, a formalization can be found in the papers [SP01], where it is also indicated how to extend this formalization to IF-predicate logic, and [SP03], which paper studies to which extent the IF-propositional games satisfy some conditions that are usual in game theory. Extensive semantic games for IF-modal logic are defined in [Tul03].

We agree with van Benthem that game-theoretic results and discussions at least shed an interesting light on logic in general, and the semantic games of imperfect information for IF-logic in particular (cf. [vB00a] and [vB05], [vB01], [vB02b], [vB03], [vB04]). In section 4.7, we show how *imperfect recall* is omnipresent in the semantic games, while situations of imperfect recall are traditionally avoided in game theory because of conceptual difficulties. In section 4.8 we will point out correspondences between game equivalence (by the so-called Thompson transformations) and logical equivalence schemes. The transformations indicate how to deal accurately with imperfect information if we want to generalize equivalence schemes from propositional or first order logic.

But first, we give the details of the model we will use to formalize semantic games in game-theoretic terms. In this part of the chapter, we do not assume familiarity with game-theoretic terminology.

4.2 Games in extensive form

Game theory deals with a very general class of strategic interactions, far more general than our abstract semantic games. Our games can be characterized as *two-player, non-cooperative, win-loss games of imperfect information.* Game theory has two main ways of representing games: in strategic form and in extensive form.

In a *strategic form*, an interaction is modeled by the players simultaneously choosing a strategy. Every player makes just one decision, at the start of the game, and immediately afterward the outcome is determined. Mathematically, a game in strategic form is an n-dimensional matrix, where n is the number of players, with the numbers of columns/rows in each dimension determined by the number of possible strategies for the corresponding player. The entries in the matrix give the respective outcomes for the players given the chosen strategies. Probably the most famous game-theoretic example, the Prisoner's dilemma, is well represented in this form (we slightly adapt [OR94, Fig. 17.1]):

I\II	Confess	Don't confess
Confess	-3\-3	0\-4
Don't confess	-4\0	-1\-1

(It represents two suspects (I,II) in one crime, who are interrogated separately. They can both either confess and cooperate, or deny and deflect. But their sentence depends on the combination of both player's choices. The number of years of imprisonment they get in all cases are indicated as negative numbers in the matrix. Analysis of this game in terms of *Nash equilibria* learns that the optimal choice for both players to make is to confess.)

In a sense, all games in strategic form are of imperfect information, because the players decide simultaneously, hence without knowing the other players' choices. (This makes the study of repeated games especially interesting for games in strategic form: repetition allows for reasoning on the other players' expected behavior.)

The *extensive form* represents games as a tree structure with parameters, which allows for a more dynamic view of the game and distinctions between perfect and imperfect information, and perfect and imperfect recall. It corresponds most directly to Hintikka's definition of the semantic games in terms of their rules, and gives us a formalization of the concept of strategy. We give an introduction to the more general extensive game representation (section 4.2). In section 4.3, we define a generalization of the IF-language, and give a thorough formalization of the corresponding semantic games. In sections 4.5–4.6 we collect several reflections on relations between the game structures and the syntax of IF-logic.

We first give the definition of a general mathematical model for games in extensive form, based on [OR94, pp. 200–203], which in its turn goes back to the model introduced by [Kuh53] as a reformulation of the model by Von Neumann and Morgenstern ([NM47]). In the light of the intended application, viz. modeling of the semantic games of GTS, we leave out the so-called *chance moves* that are included in the original model. Like [OR94], we impose the *consistency condition* on our informations sets, rather than employing an *alternative partition* as in the original model of [Kuh53]. Also, we choose to model the outcomes by a function (like in [vD91]) rather than by a preference relation (like in [OR94] and [Kuh53]), because a payoff function can be easily simplified for the representation of *win-loss* games like our semantic games.

In the definition of the model, we use the following notations for sequences:

Notation 4.2.1 (finite sequences) *For any set A ('the alphabet'), let $A^* = \bigcup_{n \in \mathbb{N}} A^n$ be the set of finite sequences of elements of A. For $\alpha = (a_1, \ldots, a_k) \in A^*$ and $a \in A$, we use the following notations:*

- $\mathtt{len}(\alpha) := k$;
- *for* $m \leq \mathtt{len}(\alpha)$: $\alpha|m := (a_1, \ldots, a_m)$;
- $\alpha \cdot a := (a_1, \ldots, a_k, a)$;
- $a \in \alpha :=$ *there exists* $j \leq \mathtt{len}(\alpha)$ *with* $a = a_j$.

Finally, $()$ denotes the empty sequence.

Definition 4.2.2 (games in extensive form) *A game of imperfect information* **in extensive form** *(without chance moves) is defined to be a 5-tuple $\Gamma = \langle H, N, P, U, r \rangle$, where*

- *H is a set of finite sequences that satisfies the following conditions:*
  - *$()$ is a member of H;*
  - *H is* **prefix closed**: *if $\alpha \in H$ and $m \leq \mathtt{len}(\alpha)$, then $\alpha|m \in H$.*

  These conditions makes it possible to regard the set H as a **tree structure**, *with $()$ as root. Notation: $A := \bigcup_{h \in H}\{a | a \in h\}$ (thus: $H \subseteq A^*$). Define for $h \in H$: $A(h) := \{a \in A | h \cdot a \in H\}$. Let $Z := \{h \in H | A(h) = \emptyset\}$, and $D := H \setminus Z$.*

  The elements of H will be called **histories**; *the elements of Z are the* **terminal** *histories, and the elements of D the non-terminal histories or* **decision points**. *A terminal history is called a* **play** *(cf. see [Kuh53]). $A(h)$ is the set of* **actions from** *h.*

- $N = \{p_1, \ldots, p_n\}$ *is a non-empty, finite set of* **players**. *We let* $n = \#N \in \mathbb{N}$ *be the number of players.* N *is often identified with the set* $\{1, \ldots, n\}$.

- $P : D \to N$. *Notation: for* $i \in N$ *we write* $D_i := P^{-1}(\{i\})$.

  P *is the* **player function***: it determines for each decision point whose turn it is.* D_i *is the set of decision points in which it is player* i*'s turn. The sets* D_i *form a partition of* D. *Note that one or more of the* D_i *might be empty.*

- $U = \bigcup_{i \in N} U_i$, *where for every* $i \in N$, U_i *is a partition of* D_i, *such that for all* $u \in U_i$ *the so-called consistency condition is satisfied: if* $h, h' \in u$ *then* $A(h) = A(h')$.

  U *is called the* **information partition**, *and the elements of* U *are called* **information sets***;* $A[u]$ *is the set of* **actions from** *information set* u. *Due to the consistency condition,* $A[u]$ *can be defined as* $A(h)$ *for some arbitrary* $h \in u$.

- $r : Z \to \mathbb{R}^N$ *is the* **payoff** *function, giving for all plays of the game the respective payoff for each player.*

  A two-player game is called **strictly competitive** *or* **zero-sum** *if for each* $h \in Z$*:* $r(h)(1) = -r(h)(2)$. *In that case the payoff function could be simplified into a function* $r' : Z \to \mathbb{R}$, *specifying at each* $h \in Z$ *the outcome for player* 1. *A two-player game is called a* **win-loss** *game if it is strictly competitive* and *after each play, the payoff for the winning player is the same:* $|r'(h)| = c$ *for some* $c \in \mathbb{R}$ *and all* $h \in Z$. *In that case the payoff function may be modeled as an 'outcome function'* $r'' : Z \to N$ *assigning a* winner *to each play.*

The information partition U makes this a model for *imperfect information* games: the intended interpretation of an information set u in U_i is that, if the actual history h is in the information set u, then player i knows she is in one of the positions in u, but she is not able to determine in which one. (Perfect information games are characterized by the property that all information sets are singletons.) When player i chooses an action in the actual situation h, she therefore has to choose one action from $A(h) = A[u]$ for all histories in u at the same time. This is reflected in the definition of strategy:

Definition 4.2.3 *A* **(pure) strategy** *for player* i *in the game* Γ *is a function* $F^i : U_i \to A$ *s.t. for every* $u \in U_i : F^i(u) \in A[u]$.

The addition of the adjective *pure* is to contrast this notion of strategy with the notion of *mixed* strategy, in which a player may include e.g. the toss of a coin to choose one of two possible actions: in mixed strategies a player puts a probability distribution over the set of her pure strategies ([OR94, definition 212.1]). But because mixed strategies play no role in our discussion, we refer to pure strategies simply as 'strategies'.

Note that a strategy always prescribes an action for *all* information sets assigned to the player concerned, even if an action taken at one information set

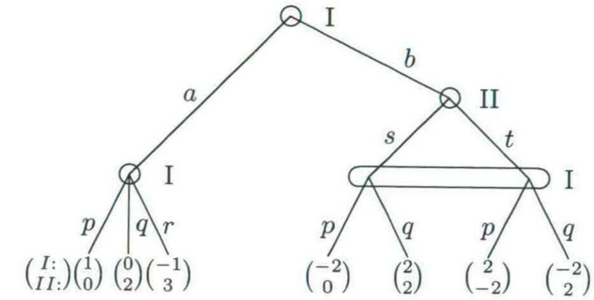


Figure 4.1: Visualization of a game in extensive form

prevents the player from ever reaching some other one. We illustrate these notions with an example:

Example 4.2.4: (a game in extensive form) Figure 4.1 gives the tree representation of the game Γ modeled by the following tuple $\langle H, N, P, \overline{U}, r\rangle$:

- The set of actions A can be read along the branches of the tree:

$$A = \{a, b, s, t, p, q, r\}.$$

  The set of histories $H = D \cup Z$ is the union of the set of non-terminal sequences in the tree (the decision points)

$$D = \{(), (a), (b), (b, s), (b, t)\},$$

  and the set of terminal histories (the plays)

$$Z = \{(a, p), (a, q), (a, r), (b, s, p), (b, s, q), (b, t, p), (b, t, q)\}.$$

- $N = \{I, II\}$

- $P : D \to N$ assigns the decision point (b) to player II, and the other decision points to player I. So, $D_{II} = \{(b)\}$ and $D_I = D - D_{II}$.

- $U = U_I \cup U_{II}$, where U_I partitions D_I into the information sets $\{()\}, \{(a)\}$, and $\{(b, s), (b, t)\}$. Note that the consistency condition holds for the non-singleton element of the partition: $A[\{(b, s), (b, t)\}] = A((b, s)) = A((b, t)) = \{p, q\}$. Because D_{II} is a singleton, U_{II} has automatically just one element: $D_{II}(= \{(b)\})$ itself.

- the payoff function $r : Z \to \mathbb{R}^N$ is depicted at the bottom of figure 4.1; it determines for example that player I loses 2 after the plays (b, s, p) and (b, t, q), but gains 2 after (b, s, q) and (b, t, p). We see that this game is not *zero-sum*: not for all plays $h \in Z$ do we have $r(h)(I) + r(h)(II) = 0$.

A strategy for player I is a function $F : U_I \to A$, so for example: the function F that assigns the action a to the information set $\{()\}$ at the root of the tree, p to the information set $\{(a)\}$, and q to the information set $\{(b,s),(b,t)\}$. Playing this strategy gives player I a payoff of 1, regardless of player II's actions.

In fact, the choice for action a in the first move prevents II from getting to make a move at all, and also guarantees that the information set $\{(b,s),(b,t)\}$ for player I is not reached. So, the action prescribed by F in $\{(b,s),(b,t)\}$ is irrelevant. However, as noted above, this 'irrelevant' action may not be omitted: a strategy for a player assigns an action to *all* information sets of that player. A reason to adhere to this, is that it allows us to speak about "the strategy F' that differs from F only in its assignment of action b to $\{()\}$". If F would not assign an action to the information set $\{(b,s),(b,t)\}$, F' would not define a full prescription for player I how to play this game.

It is not hard to verify that F is an optimal strategy, in the sense that it gives player I the highest guaranteed payoff: if she would have played b in the first information set, player I can not *ensure* that she gets the higher payoff of 2. Because she cannot distinguish between the histories (b,s) and (b,t), both the action p and the action q (prescribed by F) leave the possibility open for the lowest payoff (viz. -2). $\diamond$

We have now introduced the general model of games in extensive form. In section 4.4, we will specify semantic games in terms of this standard game model. But first, we take the opportunity to switch to a more general, symmetric language, which we will call $\mathcal{L}_{IFG}$.

4.3 The generalized language $\mathcal{L}_{IFG}$

From now on, we will use a generalization $\mathcal{L}_{IFG}$ of the languages $\mathcal{L}_{IF}$ as it was defined by Hintikka (cf. definition 2.5.1), and also of its symmetric variant $\mathcal{L}_{IFS}$, which we used in the previous chapter (cf. definition 3.4.2). We will introduce this language and some useful syntactic notions first.

We will model the semantic games for the language $\mathcal{L}_{IFG}$ (defined as $\mathbf{L_{ii}}$ in [CK99]; the subscript '**ii**' stands for *imperfect information*). This language is a natural generalization of the language of IF-logic as defined in e.g. [Hin96], and a sublanguage of the language presented by [Hod97a], as it lacks the indexed connectives (i.e. formulas of the form $\bigwedge_{i \in I} \psi_i$).

We explain how the IFG-language is more general than the IF-language. First, the IFG-language does *not* presuppose negation normal form (as the IF-language *did*). While this does not make the IFG-language more general than the IFS-language yet, the following does: the IFG-language does not contain only sentences, like both $\mathcal{L}_{IF}$ and $\mathcal{L}_{IFS}$. In IFG-formulas variables may occur freely (possibly under slashes). This will be important in the next two chapters. Finally: the IFG-language allows arbitrary sets of variables under the slashes, not just variables whose values 'were chosen by the opponent', as in IF- and IFS-sentences.

Definition 4.3.1 (The language $\mathcal{L}_{IFG}$) *Given a first order signature σ, the terms and atomic formulas of the language $\mathcal{L}^{\sigma}_{IFG}$ with equality are defined as in first order logic. The set of formulas of $\mathcal{L}^{\sigma}_{IFG}$ is defined as the least L such that*

- *each atomic formula of $\mathcal{L}^{\sigma}_{IFG}$ belongs to L;*
- *if $\varphi \in L$ then $\neg\varphi \in F$*
- *if $\varphi \in L$ and Y is a finite list of distinct variables not containing x, then $\exists x_{/Y}\varphi \in L$*
- *if $\varphi, \psi \in L$ and Y is a finite list of distinct variables, then $\varphi \vee_{/Y} \psi \in L$.*

$\forall x_{/Y}\varphi$ will be an abbreviation for $\neg\exists x_{/Y}\neg\varphi$ and $\varphi \wedge_{/Y} \psi$ will be an abbreviation for $\neg(\neg\varphi \vee_{/Y} \neg\psi)$. If $Y = \emptyset$, we omit the slash and write $\exists x\psi, \forall x\psi, \varphi \vee \psi, \varphi \wedge \psi$. We say that φ is an IFG-formula if $\varphi \in \mathcal{L}^{\sigma}_{IFG}$ for some first order signature σ. As usual, the first order signatures σ will play implicit in the phrase suitable model.

The following conventions, notations and definitions in this section, are all defined in terms of the *unabbreviated versions* of the formulas in $\mathcal{L}_{IFG}$: the formulas contain no universal quantifiers or conjunctions (these are assumed to be written out in terms of negations and existential quantifiers or disjunctions respectively). This saves us otherwise recurring case distinctions: between existential and universal quantifiers, and between disjunctions and conjunctions. This choice is justified semantically by the interpretation of the negation sign as role exchange on the one hand, and, on the other hand, the fact that universal and existential quantification (and conjunction and disjunction) correspond to the same type of move in the semantic game. If we use the negation sign to determine the player assignment, there is no need for two quantifiers, or two connectives. In our definitions, by this choice, existential quantification and disjunction indicate the (two) types of moves, while negation has a monopoly on determining the player distribution.

Usually in the literature (as in the IF-language), the opposite choice is made: both quantifiers and both connectives are used, while the negation signs occur only on atomic level (in other words, the formulas are taken to be in negation normal form). Our *later examples* will also be chosen to be abbreviated, and thereby preferably in negation normal form, because this improves their readability, and makes comparison with classical logic more direct.

Semantic games of imperfect information are defined for *IFG-sentences* in a given suitable model. To express the components of the game model in terms of these parameters, we use the following sets and notions. We will also use these definitions in the next two chapters, when we define semantics for *open* IFG-formulas as well, so we give most definitions for arbitrary *formulas* φ.

Definition 4.3.2 (subformulas) *The set of subformulas $Sub(\varphi)$ of an IFG-formula φ is the smallest set S that satisfies:*

- $\varphi \in S$

- *if* $\neg\psi \in S$*, then* $\psi \in S$
- *if* $\psi_1 \vee_{/Y} \psi_2 \in S$*, then* $\psi_1, \psi_2 \in S$
- *if* $\exists x_{/Y}\psi \in S$*, then* $\psi \in S$.

Additional remark: *This results in a* set *of formulas, which would mean that every element occurs only once (a set* $\{e, e\}$ *is the same set as* $\{e\}$*). In the logical language, syntactically identical subformulas can occur more than once in one formula, but by their position in the formula, they may have different roles, e.g. like* ψ *in the formula* $\psi \vee \neg\psi$*. We must therefore treat syntactically identical subformulas that occur more than once in* φ *as* different *subformulas, hence as* different *elements of* $Sub(\varphi)$*. For example,* $Sub(\psi \vee \neg\psi)$ *must contain the left and the right occurrences of* ψ *(and its subformulas) as distinct elements, for in the semantic games, the two different occurrences obviously define different subgames (in this case, they differ by reversal of the player assignment and designation of the winner at the atomic formulas).*

*We choose not to give a formal procedure to make the distinction between the syntactically identical formulas, but one could imagine that at connectives, we label the left conjunct/disjunct with '***l***' and the right one with '***r***', and let their subformulas inherit these labels. These labels are not part of the syntax of the subformulas, but enable us to distinguish the different occurrences.*

Notation 4.3.3 (stripping initial negation signs) *For every IFG-formula* ψ*, we use the notation* $|\psi|$ *to indicate the formula resulting from stripping all initial negation signs from* ψ*. We will frequently use the set of* **non-negated** *subformulas of* φ*:* $S(\varphi) := \{|\psi| \,|\, \psi \in Sub(\varphi)\}$.

This notation is particularly useful, because in the definition of the components, we will use the unabbreviated versions of the formulas. This means that every change of turn corresponds to a negation sign in the formula, hence the formula may contain quite a number of negations.

Definition 4.3.4 (free and bound variables) *If* ψ *is an IFG-formula, the set of free variables* $Fv(\psi)$ *and the set of bound variables* $Bv(\psi)$ *are defined inductively as usual (see 2.1.6), except for the following cases:*

($\vee$) $Fv(\psi_1 \vee_{/Y} \psi_2) := Fv(\psi_1) \cup Y \cup Fv(\psi_2)$
$Bv(\psi_1 \vee_{/Y} \psi_2) := Bv(\psi_1) \cup Bv(\psi_2)$

($\exists$) $Fv(\exists x_{/Y}\psi) := (Fv(\psi) \setminus \{x\}) \cup Y$
$Bv(\exists x_{/Y}\psi) := Bv(\psi) \cup \{x\}$

Note that $Fv(\psi)$ includes the free variables in the sets Y under the slashes at quantifications and connectives in ψ. Furthermore, note that one variable may occur both free and bound in the same formula, like x in the following (unslashed) example: $(x = y) \vee (\forall x[P(x, y)])$. In this chapter, we assume that in the IFG-sentences we work with, **no nested quantification over the same variable** occurs (thereby implying that in the subformulas of the sentences no variables occur both free and bound at the same time). Why we need this restriction, which we will call *regularity*, is demonstrated in the examples in chapter 6 (cf. the introduction to the chapter, section 6.1).

The following notations will be useful in expressing the components of the game model for a semantic game for φ in terms of the formula. We will let the histories of the general model (definition 4.2.2) correspond to pairs consisting of a subformula and a *valuation*: an assignment of values to a (finite) set of variables. This set of variables –the *domain* of the valuation– consists of the variables that are quantified higher up on the branch of ψ in the syntactic tree of the sentence φ:

Definition 4.3.5 (relative domain) *Let φ be an IFG-sentence, and let $\psi \in Sub(\varphi)$. Then the* **relative domain** *X^φ_ψ of ψ in φ is the set of variables that have scope over ψ in φ. Formally, by outside-in induction, for all $\psi \in Sub(\varphi)$:*

- $X^\varphi_\varphi := \emptyset$;
- *if* $\psi \equiv \neg\psi'$, *then* $X^\varphi_{\psi'} := X^\varphi_\psi$;
- *if* $\psi \equiv \theta \vee_{/Y} \chi$, *then* $X^\varphi_\theta := X^\varphi_\psi$ *and* $X^\varphi_\chi := X^\varphi_\psi$;
- *if* $\psi \equiv \exists x_{/Y}\psi'$, *then* $X^\varphi_{\psi'} := X^\varphi_\psi \cup \{x\}$.

It may be clarifying to note that the set of variables X^φ_ψ includes the set of free variables of ψ ($Fv(\psi) \subseteq X^\varphi_\psi$), but the two sets are not necessarily equal. This depends on the context provided by the sentence φ for ψ. For example, if ψ is the subformula $\exists y_{/x}[x = y]$ of Hodges' example

$$\varphi := \forall x \exists z \exists y_{/x}[x = y],$$

then $Fv(\psi) = \{x\}$, while $X^\varphi_\psi = \{x, z\}$. X^φ_ψ gives the set of all variables to which a value has been assigned during a semantic game for φ.

With this (relative) domain, we define for each subformula of an IFG-sentence φ, the corresponding set of valuations:

Definition 4.3.6 (valuations for subformulas) *Let φ be an IFG-sentence, $\mathfrak{A} = \langle A, \dots \rangle$ a suitable model for φ with $Dom(\mathfrak{A}) = A$, and let $\psi \in S(\varphi)$. Then $V^\varphi_\psi := A^{(X^\varphi_\psi)}$ is defined as the set of all valuations $v : X^\varphi_\psi \to A$.*
Let $V^\varphi := \bigcup_{\psi \in S(\varphi)} V^\varphi_\psi$.

To define the player distribution and the outcome function, we will use the notion of polarity of subformulas of (unabbreviated!) IFG-sentences:

Definition 4.3.7 (polarity of subformulas) *Let φ be an IFG-formula. A subformula $\psi \in S(\varphi)$ is* **positive in** *φ if it occurs under the scope of an even number of negation signs, and dually, it will be* **negative in** *φ if it occurs under the scope of an odd numbers of negations.*

4.4 Modeling semantic games in extensive form

In section 2.4, the semantic games for (IF-)first order sentences were described in terms of the *game rules*, as is done in [Hin96]. Even though this description makes it intuitively clear how the *plays* of such games go, it does not define in mathematical terms what kind of object a *strategy* is. Using the description of the semantic games in terms of their rules, we will formalize them in the mathematical model of the previous section. Similar formalizations of semantic games have been presented in [SP01] and [SP03], with a focus on an IF-propositional language.

Using the language and the syntactic notations of the previous section, we define for a given IFG-sentence φ and a suitable model $\mathfrak{A} = \langle A, \dots \rangle$ the extensive game representation $\Gamma(\mathfrak{A}, \varphi) = \langle H, N, P, r, U \rangle$ of the semantic game $G_{\mathfrak{A}}(\varphi)$. To clarify our terminology, we will illustrate the construction of the model with a concrete example:

Example 4.4.1: (A semantic game modeled in extensive form) We study the semantic game in a model $\mathfrak{A}$ with domain $\{0, 1\}$, for the IFG-sentence

$$\forall x[\exists y_{/x}(x = y) \vee \exists y_{/x}(x = y)],$$

hence, in the unabbreviated version:

$$\varphi := \neg\exists x \neg[\exists y_{/x}(x = y) \vee \exists y_{/x}(x = y)],$$

The semantic game $G_{\mathfrak{A}}(\varphi)$ is played as follows: first, Abélard (A) chooses a value from $\{0, 1\}$ as assignment for x. Then Eloïse (E) may choose whether she continues on the left or on the right disjunct. In both cases, it is then her turn again, and in both cases she may then choose an assignment from $\{0, 1\}$ for y. When she does so, she does not know which assignment Abélard chose for x (as indicated by the $_{/x}$ at the two quantifications $\exists y_{/x}$). In all cases, the game ends after that, with the evaluation of one of the atomic formulas $(x = y)$ with respect to some valuation of x and y in $\{0, 1\}$: Eloïse wins if the assignments to x and y are the same, otherwise Abélard wins. The semantic game is visualized as a game tree in figure 4.2.

We name the elements of $S(\varphi)$ (i.e. the non-negated subformulas, different occurrences distinguished) as follows:

$$\begin{aligned}
\xi &:= |\varphi| = \exists x \neg[\exists y_{/x}(x = y) \vee \exists y_{/x}(x = y)] \\
\chi &:= \exists y_{/x}(x = y) \vee \exists y_{/x}(x = y) \\
\psi_{\mathbf{l}} &:= \exists y_{/x}(x = y) \text{ (left copy)} \\
\psi_{\mathbf{r}} &:= \exists y_{/x}(x = y) \text{ (right copy)} \\
\zeta_{\mathbf{l}} &:= (x = y) \text{ (left copy)} \\
\zeta_{\mathbf{r}} &:= (x = y) \text{ (right copy)}
\end{aligned}$$

The syntactic tree of φ is pictured in figure 4.3. Some components of the game model $\Gamma(\mathfrak{A}, \varphi)$ are illustrated in figure 4.4. The construction of the components is explained in the course of this section. ◇

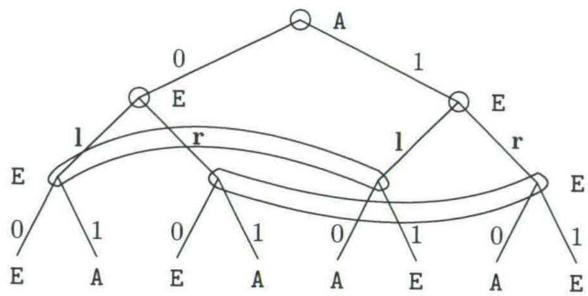


Figure 4.2: Example of a semantic game in extensive form (cf. example 4.4.1)

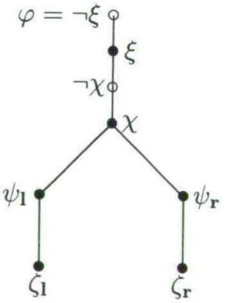


Figure 4.3: Syntactic tree of the formula (cf. example 4.4.1)

The set of histories H

The histories in a game are the possible sequences of moves from the start of the game. These sequences of moves bring us either in a position where the game prescribes one of the players to make a next move (a decision point), or to a position in which none of the players can make a move and the outcome determined (a terminal history, *play*).

In the semantic games, there are two types of moves: choice of a domain element (at a quantifier), and choice of one out of two subformulas (at a connective). The sequences of moves will therefore be sequences of elements from the set (of *actions*) $Dom(\mathfrak{A}) \cup \{\mathbf{l}, \mathbf{r}\}$. Which moves occur, and in which order, is determined by the syntax of the sentence φ. The set of histories is most easily defined by first defining for every $\psi \in S(\varphi)$ a set of finite sequences H_ψ. Note that we do this by induction on the structure of φ –from outside to inside– in a way that keeps us within $S(\varphi)$: the induction 'skips' the negation signs. It terminates at the atomic formulas.

- $H_{|\varphi|} := \{()\}$;
- Let $\psi \in S(\varphi)$ and suppose H_ψ already defined. There are three cases that can be distinguished:
  1. if ψ is atomic, induction stops;
  2. if ψ is of the form $\psi_1 \vee_{/Y} \psi_2$, then $H_{|\psi_1|} := \{h \cdot \mathbf{l} | h \in H_\psi\}$ and $H_{|\psi_2|} := \{h \cdot \mathbf{r} | h \in H_\psi\}$.
  3. if ψ is of the form $\exists x_{/Y} \psi'$, then $H_{|\psi'|} := \{h \cdot a | a \in Dom(\mathfrak{A}), h \in H_\psi\}$;

Let $H := \bigcup_{\psi \in S(\varphi)} H_\psi$, then H is the set of histories for the 5-tuple $\Gamma(\mathfrak{A}, \varphi)$ modeling the semantic game $G_{\mathfrak{A}}(\varphi)$. If $h \in H_\psi$, we say that h leads to the subformula ψ. For $h \in H$, we write ψ_h for the unique $\psi \in S(\varphi)$ such that $h \in H_\psi$.

Let D and Z be defined as the sets of non-terminal and terminal histories respectively (cf. the general case in section 4.2). We see that the set of actions is $A = Dom(\mathfrak{A}) \cup \{\mathbf{l}, \mathbf{r}\}$, and for every $h \in D$: either $A(h) = Dom(\mathfrak{A})$ (if ψ_h prompts a choice for a quantifier) or $A(h) = \{\mathbf{l}, \mathbf{r}\}$ (if ψ_h prompts a choice for a connective).

Note that $S(\varphi) = \{\psi_h | h \in H\}$. If we define $S_D := \{\psi_h | h \in D\}$ and $S_Z := \{\psi_h | h \in Z\}$, then S_Z is the set of atomic formulas in φ, and S_D is the set of non-negated, non-atomic subformulas of φ.

Every history $h \in H$ defines a valuation $v_h \in V^{\varphi}_{\psi_h}$ for the formula ψ_h:

- $v_{()} := \emptyset$;
- if v_h has been defined and $\psi_h = \psi_1 \vee_{/Y} \psi_2$, then $v_{h \cdot \mathbf{l}} = v_{h \cdot \mathbf{r}} := v_h$;
- if v_h has been defined and $\psi_h = \exists x_{/Y} \psi'$, then for every $a \in Dom(\mathfrak{A})$: $v_{h \cdot a} := v_h \cup \{(x, a)\}$.

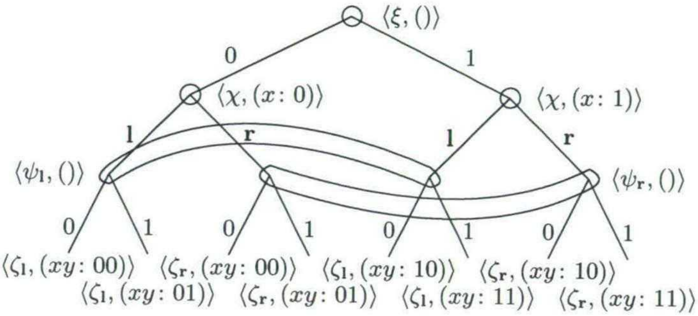


Figure 4.4: Some histories and information sets as pairs (cf. example 4.4.1)

In fact, the mapping $h \mapsto \langle \psi_h, v_h \rangle$ is a **one-one correspondence** of the set of histories H with the set $\{\langle \psi, v \rangle | \psi \in S(\varphi), v \in V_\psi^\varphi\}$, consisting of pairs of subformulas and valuations. In correspondence with the definition of semantic games (definition 2.4.1), we will sometimes call these pairs the **positions** in the game. (The duality between histories and positions corresponds to the two dual ways of representing tree structures: as a prefix-closed set of *sequences*, or as a specific type of graph, which is a set of *nodes* with a certain type of relation on them.)

Example 4.4.1 continued: (See figure 4.4.) The sequence $h := (0, \mathbf{l}, 1)$ is an example of a (terminal) history in this game. This history corresponds with the pair $\langle \zeta_\mathbf{l}, (xy\colon 01) \rangle$, consisting of the subformula $\psi_h = \zeta_\mathbf{l}$, and the valuation $v_h = (xy\colon 01)$.

The set $H_{\zeta_\mathbf{l}}$ consists of the four histories h' for which $\psi_{h'} = \zeta_\mathbf{l}$, viz. $(0, \mathbf{l}, 0)$, $(0, \mathbf{l}, 1)$, $(1, \mathbf{l}, 0)$, and $(1, \mathbf{l}, 1)$. These respectively code the valuations $(xy\colon 00)$, $(xy\colon 01)$, $(xy\colon 10)$, and $(xy\colon 11)$.

The set of players N

Semantic games are two-person games by definition, and we chose to call these players Abélard and Eloïse throughout this thesis (in accordance with a large part of the literature on logic games; see also Appendix A). Here, we use the symbols $\mathtt{A}$ and $\mathtt{E}$ to indicate them: we take the set of players N in the extensive game model to be $\{\mathtt{A}, \mathtt{E}\}$.

The player function P

We define the player function P by using the concept of polarity for subformulas in φ (cf. definition 4.3.7).

From the rules of the semantic games, it is easy to observe that all positive, non-atomic subformulas of φ will prompt a move for Eloïse, while the negative ones will prompt a move for Abélard in a semantic game for φ. In terms of the game model, the player function $P : D \to N$ is determined for each $h \in D$ by the polarity of $\psi_h \in S_D$ in φ:

$$\begin{cases} P(h) = \mathtt{E} & \text{if } \psi_h \text{ is positive} \\ P(h) = \mathtt{A} & \text{if } \psi_h \text{ is negative} \end{cases}$$

Given a decision point $h \in D$, only the subformula ψ_h defined by h is relevant for P, not the valuation v_h. In other words: P is constant on each H_ψ, and each $H_\psi \subseteq P^{-1}(i)$ for either $i = \mathtt{A}$ or $i = \mathtt{E}$. In the context of semantic games, it is natural to say that a certain *subformula* (in $S(\varphi)$) is assigned to a player, rather than the histories that lead to it. We will frequently use this convention.

Example 4.4.1 continued: (See figure 4.2.) The only negative subformula of φ is the formula ξ, so the (one) history in $H_\xi = \{()\}$ is the only decision point assigned to Abélard in this semantic game. All other decision points h are assigned to Eloïse, because their corresponding subformulas ψ_h are positive. We also say that the *subformula* ξ is assigned to Abélard, and that the other non-atomic subformulas (χ, $\psi_\mathbf{l}$ and $\psi_\mathbf{r}$) are assigned to Eloïse.

The outcome function r

Because semantic games are win-loss games, we can define r as an **outcome** function: r assigns to each terminal history $h \in Z$ a winner in N. This outcome is fixed by the polarity of ψ_h in φ, and the (classical!) evaluation of ψ_h in $\mathfrak{A}$ with respect to the valuation v_h:

$$\begin{cases} r(h) = \mathtt{E} & \text{if either: } \psi_h \text{ is positive and } \mathfrak{A} \models \psi_h[v_h] \\ & \quad\;\;\text{or: } \psi_h \text{ is negative and } \mathfrak{A} \not\models \psi_h[v_h] \\ r(h) = \mathtt{A} & \text{in the other cases} \end{cases}$$

Example 4.4.1 continued: (See figures 4.2 and 4.4.) Both atomic formulas $\zeta_\mathbf{l}$ and $\zeta_\mathbf{r}$ are positive, and satisfied by the valuations (xy: 00) and (xy: 11), and not by the other two valuations in $V^\varphi_{\zeta_\mathbf{l}} = V^\varphi_{\zeta_\mathbf{r}} = \{0,1\}^{\{x,y\}}$. This means that the histories $(0, \mathbf{l}, 0)$, $(1, \mathbf{l}, 1)$, $(0, \mathbf{r}, 0)$ and $(1, \mathbf{r}, 1)$ are winning for Eloïse, and the other terminal histories for Abélard.

The information partition U

The information sets in the game model for semantic games are induced by the slash notation in IF-logic: it is the slash notation that makes semantic games for IF-formulas into games of imperfect information. Informally, the effect of a slash at a quantifier or connective in a sentence φ, is that the player making the choice associated with it, may not use the values chosen for the variables under the slash. In other words: the player has to make the same choice as in a situation where other values had been chosen for those variables.

To express this formally, let for $\psi \in S_D$ (i.e. non-atomic $\psi \in S(\varphi)$), Y_ψ be the –possibly empty– sequence of variables occurring under the slash of the main connective or quantifier of ψ. Using this set of variables, we can define an equivalence relation on the set of valuations V^φ_ψ for ψ: if $v, v' \in V^\varphi_\psi$, then

$$v' \sim_\psi v \overset{d}{\Leftrightarrow} v'(x) = v(x) \text{ for all } x \in X^\varphi_\psi - Y_\psi$$

In other words: $v' \sim_\psi v$ if and only if v and v' coincide outside the values they assign to the variables in Y_ψ. This induces an equivalence relation $\sim_{U_\psi}$ on the set H_ψ of histories with $\psi_h = \psi$: $h \sim_{U_\psi} h'$ iff $v_h \sim_\psi v_{h'}$.

The information sets in the semantic game for φ are then the equivalence classes $u[h] = \{h' \in H_{\psi_h} | h \sim_{\psi_h} h'\}$ of the equivalence relations $\sim_{\psi_h}$, or in other words, of the equivalence relation $\sim_U$ on D, defined by:

$$h' \sim_U h \overset{d}{\Leftrightarrow} \psi_{h'} = \psi_h, \text{ and } v_{h'} \sim_{\psi_h} v_h$$

In words: two non-terminal histories are considered to be indistinguishable for the player assigned to them, if they lead to the same subformula ψ of φ *and* if the associated valuations assign the same values to the variables in $X^\varphi_\psi - Y_\psi$.

It is easy to verify that these information sets satisfy the consistency condition: for any $\psi \in S(\varphi)$, the set of actions $A(h)$ is either $Dom(\mathfrak{A})$ for all $h \in H_\psi$ (if $\psi \equiv \exists x_{/Y}\psi'$), or $\{\mathbf{l}, \mathbf{r}\}$ (if $\psi \equiv \psi_1 \vee_{/Y} \psi_2$). If two histories $h, h' \in D$ are in the same information set, they are in the same H_ψ, and hence $A(h) = A(h')$.

Of course, if ψ is 'not slashed', i.e. if $Y_\psi = \emptyset$, then the equivalence classes of the equivalence relation $\sim_\psi$ are singletons. This means that all histories leading to ψ are distinguishable. In particular, if φ is a classical first order formula, *all* information sets in the game are singletons. Semantic games for classical first order formulas are hence games of perfect information.

Concluding, we define the information partition U to be the set $\{u[h] | h \in D\}$ of equivalence classes of $\sim_U$. The partition can be divided into partitions $U_{\mathtt{E}}, U_{\mathtt{A}}$ of the sets of histories $D_{\mathtt{E}}, D_{\mathtt{A}}$ assigned to Eloïse and Abélard respectively: $U_i := \{u[h] | h \in D_i\}$. (Note that the U_i are well defined, because for $i \in \{\mathtt{A}, \mathtt{E}\}$: if $h \in D_i$, then $u[h] \subseteq H_{\psi_h} \subseteq D_i$.)

In analogy to the correspondence we had between histories h and pairs $\langle \psi_h, v_h \rangle$, we could characterize the information sets $u[h]$ by the pairs $\langle \psi_h, \overline{v}_h \rangle$, where $\overline{v}_h$ is the restriction of v_h to $X^\varphi_{\psi_h} - Y_\psi$ (i.e. a valuation to all variables of $X^\varphi_{\psi_h}$ that do not occur under the slash of the main quantifier or connective of the subformula ψ_h of φ). In order to express this correspondence formally, we define:

Definition 4.4.2 (restricted valuations, cf. definition 4.3.6) *For every $\psi \in S_D$, we define $\overline{V}^\varphi_\psi$ to be set of all valuations $\overline{v} : (X^\varphi_\psi - Y_\psi) \to Dom(\mathfrak{A})$. Furthermore, let $\overline{V}^\varphi := \bigcup_{\psi \in S(\varphi)} \overline{V}^\varphi_\psi$.*

We then have a one-one correspondence of the set U of information sets with the set: $\{\langle \psi, \overline{v} \rangle | \psi \in S_D, \overline{v} \in \overline{V}^\varphi_\psi\}$.

Example 4.4.1 continued: (See figure 4.4.) The set of histories $H_{\psi_{\mathbf{r}}}$ consists of the two histories $(0, \mathbf{r})$ and $(1, \mathbf{r})$, which correspond to the pairs $\langle \psi_{\mathbf{r}}, (x\colon 0) \rangle$ and

$\langle\psi_{\mathbf{r}},(x\colon 1)\rangle$ respectively. Because $Y_{\psi_{\mathbf{r}}}=\{x\}$, and the two valuations $(x\colon 0)$ and $(x\colon 1)$ are both the empty valuation outside of their domain $X^{\varphi}_{\psi_{\mathbf{r}}}=\{x\}=Y_{\psi_{\mathbf{r}}}$, they are in the same equivalence class of $\sim_{\psi_{\mathbf{r}}}$. So, the set $H_{\psi_{\mathbf{r}}}$ is one information set, which corresponds to the pair $\langle\psi_{\mathbf{r}},()\rangle$.

Of course, similarly, $H_{\psi_{\mathbf{l}}}$ is also one information set, which corresponds to the pair $\langle\psi_{\mathbf{l}},()\rangle$. Because Y_{ξ} and Y_{χ} are empty, the other information sets are the three singletons containing the elements of H_{ξ} and H_{χ}.

Strategies and choice functions

We claim that the 5-tuple $\Gamma(\mathfrak{A},\varphi)=\langle H,N,P,r,U\rangle$, as defined above, captures the semantic game $G_{\mathfrak{A}}(\varphi)$ as described by the game rules in section 2.4.1, with the information restrictions as described in section 2.5. The components H, P and r model the possible plays of the game with their outcomes, and the information partition U models (some aspects of) the knowledge of the players during the play.

The main focus for semantic games however, lies in the concept of strategy. We can now define it in terms of the 5-tuple, as in section 4.2. To find out how game theory models strategies, was the main aim of this modeling exercise.

Following the definition of section 4.2, a strategy for a player i in game Γ, is a function F^i assigning to each information set $u\in U_i$ an action $a\in A[u]$. For semantic games, it is natural to regard a strategy F^i for player i as a *set of* **choice functions** $f^i_\psi:\overline{V}^{\varphi}_{\psi}\to A_\psi$, one for each subformula $\psi\in S(\varphi)$ that is assigned to player i. Here A_ψ is either $Dom(\mathfrak{A})$ or $\{\mathbf{l},\mathbf{r}\}$, depending on whether ψ is a quantified formula or one joining two formulas by a main connective. Profiting from the correspondence between U and pairs $\langle\psi,\overline{v}\rangle$ with $\psi\in S_D,\overline{v}\in\overline{V}^{\varphi}_{\psi}$ (see the previous subsection), the one-function approach and the 'set of choice functions' approach can be seen to be interchangeable:

- If $F^i:U_i\to A$ is a strategy for player i we can define the corresponding choice functions $f^i_\psi:\overline{V}^{\varphi}_{\psi}\to A$ (for all ψ assigned to player i) by

$$f^i_\psi(\overline{v}):=F^i(u),$$

  where u is the information set corresponding to the pair $\langle\psi,\overline{v}\rangle$.

- If $\{f^i_\psi|\psi\in S_D,\psi$ assigned to player $i\}$ is a set of choice functions, we can form the strategy $F^i:U_i\to A$ by defining for each $h\in D_i$:

$$F^i(u[h])=f^i_{\psi_h}(\overline{v_h}).$$

In semantic games, it is more natural to think of a strategy as a set of choice functions rather than as one function. A choice function f^i_ψ prescribes a choice for player i at formula ψ on the basis of the values previously chosen for the variables in $X^{\varphi}_{\psi}-Y_\psi$.

Returning to the discussion of section 3.9 in the previous chapter, we see that the choice functions in this game-theoretic formalization correspond to **Skolem**

functions as obtained by the *inside-out* Skolemization procedure. To obtain the Skolem functions of Hintikka's approach, we would have to make the implicit independences introduced by the *slashing convention* explicit. (Cf. the quote from [Hin96] on page 44 in section 3.9, introducing the slashing convention. Remark that the quote starts with a reference to the game-theoretic notion of *information set*, while otherwise Hintikka seems to carefully avoid game-theoretic terminology. Intentionally so, as we understand from the paragraph we quoted from [Hin02a], on page 50.) This means that the main quantifiers and connectives of the positive subformulas ψ of a sentence φ will get extra variables under the slash (viz. those variables in X_ψ^φ that are quantified by the main quantifier of another positive subformula).

Example 4.4.1 continued: We give a winning strategy for Eloïse in the game $G_{\mathfrak{A}}(\varphi)$. The 'trick' behind it, is that she will use two different choice functions for the subformulas $\psi_{\mathbf{l}}$ and $\psi_{\mathbf{r}}$, even though these subformulas are syntactically identical. In the semantic game, they each define their own subgame, and Eloïse is of course allowed to choose a different plan to play in the one or the other, and coordinate this with an earlier decision which of the two subgames to play.

A winning strategy consists of the following choice functions:

- $f_X : \{0,1\}^{\{x\}} \to \{\mathbf{l}, \mathbf{r}\}$ assigns $\mathbf{l}$ to the valuation $(x\colon 0)$, and $\mathbf{r}$ to the valuation $(x\colon 1)$.
- $f_{\psi_{\mathbf{l}}} : \{()\} \to \{0,1\}$ assigns 0 (to the empty valuation).
- $f_{\psi_{\mathbf{r}}} : \{()\} \to \{0,1\}$ assigns 1 (to the empty valuation).

This strategy is *winning*, because it lets the play end in either the terminal history $(0, \mathbf{l}, 0)$, or the terminal history $(1, \mathbf{r}, 1)$ (depending on Abélard's initial move). Both histories are winning for Eloïse, so these three choice functions define a winning strategy for Eloïse.

Note that Eloïse would not have a winning strategy in the game for the formula $\varphi' = \forall x[\exists y_{/x}(x = y)]$, in the same 2-element model. This example indicates that traditional logical laws (from classical first order logic) may not always be transferable to our logic 'with imperfect information'. The semantic game for a formula $\forall x[\psi]$ (in a given suitable model $\mathfrak{A}$) is really different from the game for $\forall x[\psi \vee \psi]$ (in the same model): in the second case there are more decision points, where different actions may be taken (this is one of the observations in [Jan02]; see also [Jan97][p. 182]).

Note also that Eloïse having a winning strategy here depends on the fact that the model has two elements. In a model with three elements, she would, with a similar 'trick', have a winning strategy for the formula $\forall x[\psi_1 \vee (\psi_2 \vee \psi_3)]$, where for $k = 1, 2, 3$: $\psi_k = \exists y_{/x}(x = y)$, etc. By 'copying' subformulas, we add moves, and this can well be significant. Adding (superfluous) moves is one of the topics we discuss later, when we study the Thompson transformations on extensive games (section 4.8).

4.5 Some reflections on the extensive model

Let us now make some observations about the semantic games inspired by their game-theoretic representation. Throughout this section, we let φ be an IF(G)-sentence, and $\mathfrak{A}$ a model suitable for φ. With $G_{\mathfrak{A}}(\varphi)$ we mean the semantic game for φ in $\mathfrak{A}$, and with $\Gamma(\mathfrak{A}, \varphi)$ its game-theoretic representation.

Negation:
In the definition of semantic games in terms of their rules (see section 2.4.1), negation appears to be a dynamic element in the game: during the game, the roles of the two players are reversed. But if we compare the syntactic tree of the formula for which we play the semantic game with the game tree, we see that negation disappears in the latter. Instead, the 'meaning' of the negation signs is completely absorbed in both the player function and the outcome function (where we used the polarity of non-negated subformulas, which is in turn completely and exclusively determined by the occurrence of the negation signs in the sentence).

From a logical perspective, we can observe that **double negation cancels** naturally: adding any sequence of negations of even length anywhere in the IFG-sentence φ, will leave the 5-tuple of $\Gamma(\mathfrak{A}, \varphi)$ as defined in the previous section unchanged. This means that for example $\neg\neg\varphi$ cannot be distinguished game-theoretically from φ.

Hence, the sentence φ and a sentence φ' resulting from adding a sequence of negations of even length, are (strongly) equivalent in GTS in the sense that for *both* players holds: there exists a winning strategy for that player in $G_{\mathfrak{A}}(\varphi)$ iff there exists a winning strategy for that player in $G_{\mathfrak{A}}(\varphi')$. Even stronger: a winning strategy for a player in $G_{\mathfrak{A}}(\varphi)$ *is* a winning strategy in $G_{\mathfrak{A}}(\varphi')$, and vice versa. Because in the IFG-language, $\forall x_{/Y}$ is an abbreviation of $\neg\exists x_{/Y}\neg$ by definition, this implies that De Morgan's laws hold (e.g. $\neg\forall x\psi = \neg\neg\exists x\psi\neg \equiv \exists x\neg\psi$). They do so in the strongest possible sense (as we already noted after lemma 3.4.3).

It follows that every sentence φ has a strongly equivalent negation normal form φ' (using the abbreviations $\forall$ and $\wedge$ of course). One could say that the negation normal form of a formula is the form in which the syntax most directly represents the structure of the game tree.

Characteristic structure of the game tree:
The semantic games for IFG-sentences have a rather specific game tree: at every decision point, the tree either branches twice, or with the cardinality α of $Dom(\mathfrak{A})$ (which may be infinitely, even uncountably many times). Nodes in the game tree that correspond with nodes in the syntactic tree that do not branch, i.e. histories h that lead to quantified subformulas, branch α times. All subtrees in the game tree with one of the α histories $(h \cdot a)$ as root, have the same structure (as do all the subtrees defined by the successors $(h' \cdot a)$ of the histories h' with $\psi_{h'} = \psi_h$). Nodes in the game tree that correspond to nodes in the syntactic tree that branch in two, i.e. histories h that lead to connective-subformulas, also branch in two. The two subtrees that have either $(h \cdot \mathbf{l})$ or $(h \cdot \mathbf{r})$ as root, can have a different structure. However, for all histories h' with $\psi_{h'} = \psi_h$, the subtrees defined by

$(h' \cdot \mathbf{l})$ is of the same structure as the one defined by $(h \cdot \mathbf{l})$, and similarly for $(h' \cdot \mathbf{r})$ and $(h \cdot \mathbf{r})$.

The structure of the game tree for the semantic game $G_{\mathfrak{A}}(\varphi)$ is mostly determined by the syntax (syntactic tree) of the sentence. The model also influences the game tree by determining the number of branches at histories (nodes) corresponding with quantified subformulas. The model is the *main* parameter in determining the outcome function, as it determines the interpretation of the predicate symbols and thereby the evaluation of the atomic formulas. The syntax of the sentence is still relevant for the determination of the polarity of the atomic formulas corresponding to the terminal histories.

It does not seem easy to give a natural game-theoretic classification for this type of game structures. The specific structure of the semantic games will be one of the obstacles we encounter when we try to apply the Thompson transformations, in particular *coalescence* (in section 4.8.4).

Structure of the information sets:
In the model $\Gamma(\mathfrak{A}, \varphi)$ for the semantic game $G_{\mathfrak{A}}(\varphi)$, the information sets are induced by the equivalence relations $\sim_\psi$ on the sets of valuations V_ψ^φ for the non-negated, non-atomic subformulas ψ of φ. In turn, the equivalence relations $\sim_\psi$ are determined by the sets of variables Y_ψ occurring under the slash of the main quantifier or connective of ψ.

This results in a very specific structure of the information sets: for each non-atomic $\psi \in S(\varphi)$, the equivalence classes of the $\sim_\psi$ divide the sets V_ψ^φ in equivalence classes of the same size. To be specific: the number of equivalence classes equals the cardinality of the set of restricted valuations $(Dom(\mathfrak{A}))^{X_\psi^\varphi - Y_\psi}$, while each equivalence class is of the cardinality of the set $(Dom(\mathfrak{A}))^{Y_\psi}$. All histories in one set V_ψ^φ have equal length, as their length is determined by the depth of ψ in the syntactic tree of φ. Therefore, the so-called **Von Neumann-Morgenstern condition**, viz. that all histories in one information set have equal length (cf. [SP03, p. 29]), is automatically satisfied in semantic games for IFG-sentences.

This has to do with the fact that the only type of moves that occur under the slash, are the quantifier moves. The only sort of imperfect information occurring in the semantic games is lack of information about the assignments to some variables, not about which subformula one is playing (and thereby evaluating). The latter would be the imperfect information associated with connective-moves. A natural question to ask is what happens to the model, or more specifically, what kind of information sets we get when we would allow connective-moves under the slash. We will investigate this in the next section.

In the game-theoretic model, the treatment of the two players is naturally symmetric, which corresponds nicely to the point of taking both truth and falsity into account (cf. the previous chapter). We end this section suggesting a slightly different way of modeling semantic games in case one wants to focus on truth only.

Truth only: semantic games as decision problems
If one chooses to disregard falsity in terms of Eloïse's opponent having a winning

strategy, like Hintikka in general (see the previous chapter) and Väänänen in the perfect information semantic game (see [Vää02]: in this higher-level version, the moves in the game are choices for choice functions), it may be more natural to regard the semantic games as so-called *decision problems.* These are games for just one player ('Myself'), playing against 'Nature' (names Hintikka also uses for the players in his earlier work on GTS). The non-terminal histories assigned to Nature, are modeled by *chance moves.* As we did not include these in the model of section 4.2, we give a brief informal description of them here.

Chance moves are modeled as non-terminal histories that are assigned to a *non-personal* player (player 0, or 'Nature') The non-personal player does not play with some (unknown) strategy, but according to a probability distribution which is known to all players (it is a parameter of the game). For instance, flipping a coin can be modeled as a chance move, with probability distribution $(\frac{1}{2}, \frac{1}{2})$ for the actions `head` and `tail` (if it's a fair coin).

We can transform a semantic game into a decision problem, by modeling the non-terminal histories associated with the negative subformulas (i.e. Abélard's decision points) as chance moves. It then becomes a game with Eloïse as single player. We give all chance moves some probability distribution that gives all actions a probability greater than zero (we have to note that we may have to deal with uncountably many alternatives, e.g. if we play on the set $\mathbb{R}$). Then there exists a (pure) winning strategy for Eloïse in this decision game if and only if a winning strategy exists for her in the semantic game. If truth of a sentence is defined as Eloïse having a winning strategy in the decision problem, this coincides with the definition of truth in terms of the semantic game.

4.6 Independence of connectives?

Like quantifiers, connectives also have a *scope*, which is nicely visible in the so-called Polish notation $\vee(\psi_1, \psi_2)$ for $\psi_1 \vee \psi_2$. This means that, in principle, we might use the slash notation also to remove quantifiers and connectives in the disjuncts ψ_i from the scope of this disjunction. For example, the slashes in the sentence $\varphi = \exists x_{/\wedge} P(x) \wedge \exists x_{/\wedge} R(x)$ can be interpreted: after Abélard's choice for either the left or right conjunct, Eloïse has to choose a value for x, but she does not know whether she is choosing it to satisfy $P(x)$ or $R(x)$.

Neither the IF-language defined by Hintikka in [Hin96], nor the IFG-language we defined, include this possibility. In this section we demonstrate by a list of examples, adding to the discussion of [SP03] for the IF-propositional case, that in many cases slashing for connectives cannot be interpreted game-theoretically (i.e. in terms of information sets).

As preliminary remark, we would like to make clear that we do not consider the possible ambiguity of reference a problem. Many connectives of the same kind may occur in the formula (especially if we write the formula in its unabbreviated version, when only disjunctions occur). But this ambiguity could easily be solved by applying some kind of labeling to the different occurrences of the respective connectives.

In the formula φ we gave as example in the first paragraph of this section, the slashes at the two quantifiers indeed make sense. But this is not always the case, as will be demonstrated in the following series of variations of φ. The notion of information set as defined in the standard extensive game model, helps explain why in many situations putting a connective under a slash will not alter the interpretation of the formula in game-theoretical semantics.

- Consider $\varphi_1 = \exists x_{/\wedge} P(x) \wedge R$. After Abélard's first move, Eloïse either has to choose a value for x (if Abélard chose 'left') or do nothing (if Abélard chose 'right'). The conjunction under the slash would, extending the interpretation for variables under a slash, be interpreted as 'when Eloïse chooses a value for x, she does not know whether Abélard chose left or right'. But the mere fact that she is prompted to make a move, removes this imposed ignorance. The conjunction under the slash therefore does not alter the game.

  In terms of information sets: the histories $(\mathbf{l})$ and $(\mathbf{r})$, leading to the subformulas $\exists x_{/\wedge}\psi$ and R respectively, can never be in the same information set, because information sets never contain terminal histories (and $(\mathbf{r})$ is terminal). So, Eloïse's information set is the singleton $\{(\mathbf{l})\}$.

- Now let $\varphi_2 = \exists x_{/\wedge} P(x) \wedge \exists x R(x)$. We give a tentative interpretation of this formula. After Abélard's choice for left or right, Eloïse has to pick a value for x in either case. If Abélard chose right, she knows that this value has to satisfy $R(x)$, otherwise she does not know whether it has to satisfy $P(x)$ or $R(x)$. But in the latter case, the fact that she is ignorant tells her that Abélard must have chosen left. We assume here that she knows the structure of the game (by knowing the syntax of the sentence φ_2). It is then like a game in which the opponent announces before moving: "After my move, I will not tell you what my move was if it was 'left', but I will tell you what my move was if it was 'right'." (Note that if we could have chosen different variables in the two conjuncts, e.g. $\exists x_{/\wedge} P(x) \wedge \exists y R(y)$. In that case, the name of the variable could have given Eloïse an extra *signal* that Abélard chose right. But we do not think this is essential, and it does not come back in the description of the situation in terms of information sets.)

  In terms of information sets: because the information sets form a partition, it follows from the fact that $\{(\mathbf{r})\}$ is a singleton information set, that $\{(\mathbf{l})\}$ must also be a singleton (for these two histories are the only decision points assigned to Eloïse).

- Similarly, consider $\varphi_3 = \exists x_{/\wedge} P(x) \wedge \forall x_{/\wedge} R(x)$. Like with φ_1, the fact that Eloïse is prompted to make a move, tells her that Abélard must have chosen left in the first move. Similarly, even if we assume that Abélard forgets his first move, he can reconstruct that it must have been a choice for right if he is prompted to make another move.

  In terms of information sets: an information set never contains two histories that are assigned to different players (for the information partition $U =$

$U_{\mathtt{E}} \cup U_{\mathtt{A}}$ is defined to partition the sets $D_i = P^{-1}(\{i\})$ of decision points assigned to player i). So, Eloïse has a singleton information set $\{(\mathbf{l})\}$, and Abelard a singleton information set $\{(\mathbf{r})\}$.

- Also, consider $\varphi_4 = \exists x_{/\wedge} P(x) \wedge (R_1 \vee_{/\wedge} R_2)$. After both possible choices of Abélard, it's Eloïse's turn to make a move. But in order to make the move, she has to know whether to choose an action from $Dom(\mathfrak{A})$, or from $\{\mathbf{l}, \mathbf{r}\}$. And knowing this, implies knowing in which of both possible histories she is.

  In terms of information sets: because the two histories $(\mathbf{l})$ and $(\mathbf{r})$ define different sets of actions $(A((\mathbf{l})) \neq A((\mathbf{r})))$, they cannot be in the same information set without violating the the consistency condition we imposed.

We also look at some more complicated examples, where a slash for the same connective appears at different depths within the formula:

- First, let $\xi_1 = \exists x_{/\wedge} P(x) \wedge \exists y \exists x_{/\wedge} R(x, y)$. We try to interpret the slashes for this formula. When the actual history is leading to the subformula $\exists x_{/\wedge} R(x, y)$, she can distinguish that history from the other histories leading to that subformula, by the value she previously assigned to y. Does the fact that she gets to make a second move make her aware of the fact that Abélard chose to go 'right' at the initial connective move in the game? If the actual history is leading to the subformula $\exists x_{/\wedge} P(x)$, does she know she has not assigned a value to the variable y? However, in general, we do not assume that the players always know the number of previous moves: information sets may in general contain histories of different length. Imposing

  In terms of information sets, we get only one reasonable interpretation of ξ_1. Note that the history $(\mathbf{r})$ leading to the subformula $\exists y \exists x_{/\wedge} R(x, y)$ forms a singleton information set. For the same reason as in the example φ_2 above, the history $(\mathbf{l})$ is in a different information set. Because all histories $(\mathbf{r}, a)$ (with $a \in Dom(\mathfrak{A})$) are distinguishable for Eloïse, they each are in different information sets. Hence, if the history $(\mathbf{l})$ would not form a singleton information set, it would form an information set together with just one history $(\mathbf{r}, a_0)$, while all histories $(\mathbf{r}, a)$ with $a \neq a_0$ would form singleton information sets. Even though this situation is technically possible (histories in one information set are not required to be of the same length), the formula ξ_1 does not specify a_0 in any way and seems to be arbitrary. The alternative of taking $\{(\mathbf{l})\}$, as all other histories for Eloïse, to be a singleton information set is therefore the most reasonable choice.

- If we consider $\xi_2 = \exists x_{/\wedge} P(x) \wedge \exists y \exists x_{/\wedge, y} R(x, y)$, we see two possible interpretations: either Eloïse can, or she cannot distinguish between the choice for x in the subformula $\exists x_{/\wedge} P(x)$ and in the subformula $\exists x_{/\wedge, y} R(x, y)$. We could argue for the first option on the basis of the assumption that she does remember whether or not she made a previous move. But as noted in the previous example: we do not assume in general that players count the moves. So, let's say she cannot distinguish between the situations.

In that case, Eloïse has two information sets: $\{(\mathbf{r})\}$ and $\{(\mathbf{l})\} \cup \{(\mathbf{r}, a) | a \in Dom(\mathfrak{A})\}$. This give us an information set that contains histories histories of unequal length which also means that they don't all lead to the same subformula of ξ_2. This is technically possible if we choose not to impose the Von Neumann-Morgenstern condition. The consistency condition is not violated because in this case both subformulas prompt the same kind of move for Eloïse.

- The last variant we discuss, is $\xi_3 = \exists x_{/\wedge} P(x) \wedge \exists y_{/\wedge} \exists x_{/y} R(x, y)$. This one does have an intuitively clear interpretation: after Abélard's move, Eloïse gets to choose a domain element, but she does not know whether it will serve as value for x in the left conjunct, or as value for y in the right conjunct. If she is prompted to do a second move, she now knows they are playing the right conjunct. She then however has forgotten her previous choice for y. So, she chooses a value for x without using the value for y she chose in the previous move.

  In terms of information sets: Eloïse has two information sets, viz. $\{(\mathbf{l}), (\mathbf{r})\}$ and $\{(\mathbf{r}, a) | a \in Dom(\mathfrak{A})\}$.

From the last two examples (ξ_2 and ξ_3), we do observe that in some cases slashing for connectives can make sense, and that it extends the class of game trees for semantic games. Note that if histories leading to different subformulas are in the same information set (as was the situation with ξ_2), we need to review our formalization of strategies as sets of choice functions f_ψ, one for each $\psi \in S(\varphi)$ (see the paragraph on strategies on page 65). This illustrates that slashing for connectives results in a different structure of the information sets, and in that sense is of a really different nature from slashing for quantifications.

From the complete list of examples, we conclude that including the introduction of independence of connectives can lead to interpretation difficulties, and that in many situations a connective under a slash does not add new imperfect information (in terms of bigger information sets). We therefore choose not to extend the language $\mathcal{L}_{IFG}$ with this possibility in the remainder of this thesis.

But we do want to mention that it seems possible to model the semantic games for formulas with slashed-out connectives by information sets, if a slashed-out conjunction (or disjunction) occurs at the same quantifier in both conjuncts (or disjuncts), or at the same connective in both conjuncts (or disjuncts). The formula $\exists x_{/\wedge} P(x) \wedge \exists x_{/\wedge} R(x)$ we started this section with, is an example of the first kind, and any formula with the four-place connective W from [SP01], $W(\varphi, \psi, \chi, \xi) = (\varphi \vee_{/\wedge} \psi) \wedge (\chi \vee_{/\wedge} \xi)$, is of the second kind.

Example 5.1 of [Hod97a] uses a formula similar to $\exists x_{/\wedge} P(x) \wedge \exists x_{/\wedge} R(x)$ as an example showing that an analysis of these formulas in terms of Skolem functions (rather than by use of information sets) is more problematic. This is due to the fact that the connectives do not appear as arguments of the Skolem functions, which makes it unclear how to incorporate the independence of the connectives under the slash in the Skolemfunctions.

In the two examples mentioned above, we can treat the connectives as bounded quantifications, e.g. read $\exists x_{/\wedge}\varphi_1(x) \wedge \exists x_{/\wedge}\varphi_2(x)$ as $\wedge_{i=1,2}\exists x_{/i}\varphi_i(x)$, or

$$(\psi_{11} \vee_{/\wedge} \psi_{12}) \wedge (\psi_{21} \vee_{/\wedge} \psi_{22})$$

as

$$\wedge_{i=1,2}(\vee_{j=1,2})_{/i}\psi_{ij}.$$

If we extend our valuations with assignments to the indexes of these new, finite, quantifications, slashing for connectives becomes slashing for the indexes (just like the slashes for the quantified variables). It then also becomes possible to include these indices in the Skolem-functions. Indeed, this is part of the language used in [Hod97a] (which is otherwise the same as our language $\mathcal{L}_{IFG}$). We remark however, that formulas like $\varphi_1, \varphi_2, \varphi_i, \psi_{ij}$, are *not* part of the language, but part of the metalanguage. The formula $(\vee_{i=1,2})\varphi_i$ does not mean anything, until we specify in the metalanguage which formulas (within the language) φ_1 and φ_2 in fact are. Only if our language is assumed to contain relation symbols with indexes, we can build up formulas with indexed connectives within the language itself, e.g. $(\wedge_{i=1,2})\exists x_{/j}R_i(x)$. Note that in that case, the formulas always have the symmetric structure needed for meaningful slashing for connectives (which can be recognized in the game tree by the fact that all subtrees under the indexed connective are the same, except for the values of the outcome function at the leaves).

With indexed connectives, we generalize the syntax to include n-ary connectives for arbitrary n, and not just binary connectives. This influences the structure of the corresponding game trees, in that the number of branches at decision points can then be either n for arbitrary n, or the cardinality α of $Dom(\mathfrak{A})$. We will refer to this in section 4.8.4. (Note that, as a side-effect, the indices automatically prevent problems with ambiguity of reference, if we want to put a conjunction under a slash while more that one conjunction occurs in one formula.)

These observations conclude the first part of the chapter, in which we have given a step by step translation of semantic games into the game-theoretic framework. The rest of the chapter will have a different character. Having embedded semantic game theory, we attempt to use *off the shelf* results from game theory to obtain logical insights and/or results. In this attempt, we will be much more schematic and less formal. We hope our informal (and visual) accounts of the definitions and results give enough intuition to follow our arguments.

4.7 Imperfect recall in IFG-semantic games

Perfect recall is, informally, the property of a game that "each player is allowed by the rules of the game to remember everything he knew at previous moves and all of his choices at those moves" ([Kuh53]). A commonly used reformulation is that of [Sel75] ([Kuh97, p. 319], also used in [Bno04]).

Since the introduction of the distinction between perfect and imperfect recall for extensive games by Kuhn, "traditional game theory has excluded games with imperfect recall from its scope" ([PR97, p. 4]). The study of games of imperfect recall was given new interest by the paper cited here, which gives several different possible interpretations for information sets and strategies for decision problems (i.e. one-player games) of imperfect recall, and shows how different interpretations allow for different kinds of analysis.

We make some short remarks on how imperfect recall occurs in semantic games for IFG-sentences. It is easy to see that semantic games may be of imperfect recall: e.g. in the formula $\exists x \exists y_{/x}[x = y]$, Eloïse does not know the value for x she chose when she has to choose a value for y. (Note that this does not prevent her from having a winning strategy, cf. the discussion in section 3.9, p. 45.) In fact, in IF-logic as defined by Hintikka, *most* semantic games are of imperfect recall because the slashing convention makes Eloïse forget all her previous moves (cf. section 3.9).

In [Bno04], the notion of perfect recall is proved to be axiomatized by two independent properties:

- Action Recall (AR): in every decision point, the players know what they *did* before. (Formal definition: [Bno04, p. 248])
- Memory of previous Knowledge (MK): in every decision point, the players know what they *knew* before. (Formal definition: [Bno04, p. 244], using KM for "Knowledge Memory" instead of MK)

Imperfect recall can be caused by a failure of either the first or the latter, or by both at the same time.

In semantic games for IF(G)-sentences, imperfect recall occurs in both forms. A semantic game for the formula $\exists x \exists y_{/x}[x = y]$ is an example of a failure of Action Recall (AR), without failure of Memory of previous Knowledge (MK). In fact, the semantic game for this formula in a two-element model is identical to the example given in [Bno04, Fig. 2, p. 240]. The slashing convention makes AR fail in the semantic games for most IF-sentences.

Perhaps more surprisingly, Hodges' formula $\forall x \exists z \exists y_{/x}[x = y]$ turns out to define semantic games of imperfect recall as well.[1] It gives rise to failure of Memory of previous Knowledge, without failure of AR. The failure of MK is due to the fact that, choosing a value for y, Eloïse does not know the value of x anymore, while she knew it before. (By the fact that AR does not fail here, she still has a winning strategy: she does remember her own previous move.) The failure of MK can be recognized in the structure of the information sets in the game tree for Hodges' formula, which is the left game tree in figure 4.7 on page 78. The formalization of MK says that if an initial part of a history h is in an information set u assigned to the same player, then all histories in the information set containing h have an initial part in u. In figure 4.7 we see that this condition fails for both information sets associated with the subformula $\exists y_{/x}[x = y]$. (For the same reason, our example 4.4.1 is also a game with failure of MK for Eloïse, as we can see in figure 4.2 on page 60.)

[1] We realized this when studying the notion of perfect recall in [Bno04]; afterwards we found that this observation is also made in [vB05].

An interesting result in [Bno04] is Proposition 6: MK rules out so-called absent-mindedness. Absent-mindedness is typical of the central example in the article [PR97]: the absent-minded driver (Fig. 1, loc. cit.). It is characterized by the fact that a history and some initial part of it are in the same information set (cf. [Bno04, p. 246], definition 4).

As we saw previously, by Hodges' formula, MK can fail in semantic games for IFG-logic. So, we may wonder whether there are semantic games with absent-mindedness. But this possibility is easily ruled out by the fact that all histories in one information set of a semantic game for an IFG-sentence have equal length. Even if we would allow connectives under the slash, which would allow for histories of different length to be in one information set (cf. section 4.6), one history in an information set could never be an initial part of another history in the same information set: if they are of different length, then they must differ at the position corresponding to the connective. (This demonstrates that the converse of Proposition 6 of [Bno04] does not hold: failure of MK does not imply Absent-mindedness.)

Note that we avoid imperfect recall in semantic games by switching to the team interpretation mentioned in section 3.9, but this gives us the task of conceptualizing game-theoretically what it means to be on the same team.

4.8 Thompson transformations for IFG-logic?

The Thompson transformations, named after F.B. Thompson who introduced them in his paper *Equivalence of games in extensive form* [Tho52], are four transformations which carry a game of imperfect information in extensive form into an *equivalent* one (in a game-theoretic sense, which we explain below). In this section, we relate these transformations on extensive games to syntactic operations on IF-formulas, hence look for correspondences between game-theoretic and logical equivalence schemes.

The game-theoretic sense of equivalence connected with the Thompson transformations, is defined in terms of preservation of the *reduced normal form*. The reduced normal form of a game in extensive form is relative to a given order of the set of players, and has the following structural characteristics (cf. [Tho52, Definition 12]):

- each player has just one information set;
- each player has one move in every play (i.e. each branch of the tree intersects all information sets);
- the tree is 'minimal' in the sense that no two different actions from one information set lead to isomorphic subtrees.

A game and the result of a Thompson transformation applied to it, both have the same reduced normal form. In fact, the main feature of these transformations is that they can transform any game *into* reduced normal form (cf. [Tho52, Theorem 19]).

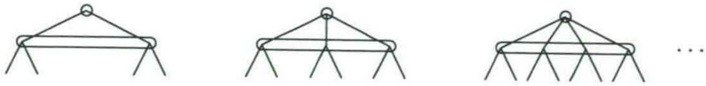


Figure 4.5: different semantic games for $\forall x[P(x) \vee_{/x} R(x)]$

Semantic games for IF-logic can be formalized as extensive games of imperfect information, as we demonstrated in section 4.4. Hence, as suggested in [vB01], a natural question to ask is to which extent the Thompson transformations define logical equivalence schemes for IF-logic. It turns out not to be hard to find logical equivalence schemes that corresponded with the respective transformations. Some of the equivalence schemes we state in this section, are results that will be formally proved in chapter 6, using the semantics of [CK99] for open IFG-formulas. The logical equivalence in the schemes is *strong* equivalence as we used it in the previous chapter (cf. definition 3.3.2).

We warn the reader in advance that the correspondence between the game transformations and the syntactic operations is not a direct one, in two ways. First, we have to realize that a syntactic operation on a subformula may correspond to many applications of a transformation: it has to be applied in *all subtrees* of the game tree corresponding to that subformula. And second: strictly speaking, the transformations imply equivalence only for one given game $\Gamma(\mathfrak{A}, \varphi)$, i.e. the semantic game for a sentence φ in *one* given model $\mathfrak{A}$, and its transformation. On the other hand, *logical* equivalence is a statement about equivalence in *all* models. Even though different models $\mathfrak{A}$ can give different game structures and outcome functions for the same formula (cf. figure 4.5, which pictures the trees for models with two, three and four elements respectively), the similarity of the game structures in different models allow us regard a Thompson transformation as operating on the *collection* of all game trees $\Gamma(\mathfrak{A}, \varphi)$ for some given φ.

We now consider the four transformations one by one. It would go too far in this context to give complete formal definitions of the transformations (and in fact, the original formalizations in [Tho52] are unclear about some details). Therefore, we copied the pictures given in Thompson's original paper in order to introduce the transformation; we reproduced them here as figures 4.6, 4.8, 4.9 and 4.10. To avoid confusion, we stress that these pictures do *not* correspond to the IF-examples we give in connection with each transformation: they illustrate the transformation *in general* for games in extensive form of imperfect information.

4.8.1 Inflation-deflation

The principle of inflation-deflation is demonstrated by figure 4.6: we see that two information sets for player I are combined. It is based on the idea that a player may be able to infer some information on the actual history on the basis of a previous move by herself. Whether this is possible, depends on the structure of the information partition. More specifically, two information sets for one player may be combined, if any history in one information set can be distinguished from any

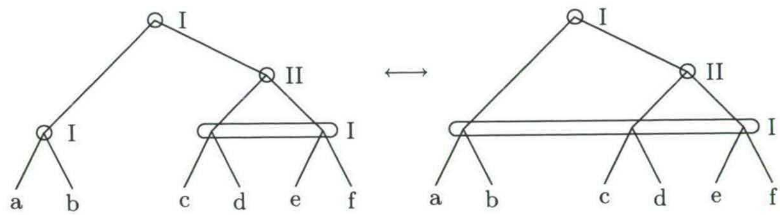


Figure 4.6: Inflation-Deflation

history in the other, on the basis of the actions taken at *one* previous information set of the same player. (In particular, a precondition for combining two information sets by this transformation, is the existence of one information set of the same player containing predecessors of all histories in the two information sets.) After the combination, the player should still be able to distinguish the two original information sets, e.g. by smartly thinking out her strategy in advance.

A clear instance of an IFG-equivalence in which this principle is reflected, is the following (cf. [vB00b, 12.3.4]):

$$\exists x \exists y[x = y] \equiv \exists x \exists y_{/x}[x = y]$$

Of course, the left sentence is true in all models. For the right one, this is not so obvious: on the basis of the game tree one would expect this formula only to be true if the model contained just one element (just as the formula $\forall x \exists y_{/x}[x = y]$). But Eloïse does have a winning strategy: she may decide in advance to play 'always a' in the first move, and also 'always a' in the second move (for some a in the domain of $\mathfrak{A}$).

In chapter 6, we will prove more generally (see lemma 6.4.7): if $\varphi, \varphi_1, \varphi_2$ are IFG-formulas and $Z \subseteq Y$, then

$$\exists x_{/Y} \exists y_{/Z} \varphi \quad \equiv \quad \exists x_{/Y} \exists y_{/Zx} \varphi \tag{4.1}$$

$$\exists x_{/Y}(\varphi_1 \vee_{/Z} \varphi_2) \quad \equiv \quad \exists x_{/Y}(\varphi_1 \vee_{/Zx} \varphi_2) \tag{4.2}$$

The condition that $Z \subseteq Y$, reflects the applicability of the transformation. At her second move, Eloïse can distinguish between the different values for x, despite the x occurring under the slash: if the choice for x was based on the values of the variables outside Y, and if $Z \subseteq Y$, then Eloïse can 'recalculate' this chosen value for x on the basis of the values of the variables outside $Z \cup \{x\}$.

This transformation can be seen in relation to the fact that in many cases the extra provision of Hintikka that the players are assumed to 'forget' all their previous moves (the slashing convention, see section 3.9), does not make a difference. In Hodges' example (3.10), the slashing convention did make a difference (the right IFG-formula reflects the reading of the left IFG-formula under the slashing convention):

$$\forall x \exists z \exists y_{/x}[x = y] \not\equiv \forall x \exists z \exists y_{/x,z}[x = y]$$

This means the transformation should not be applicable in this case. And indeed, it violates the precondition of the transformation of Inflation: there is not *one*

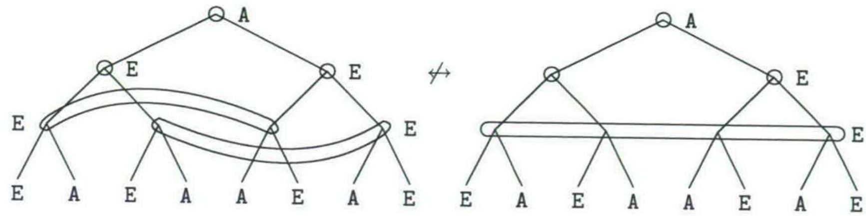


Figure 4.7: Inflation-Deflation not applicable for $\forall x \exists z \exists y_{/x}[x = y]$

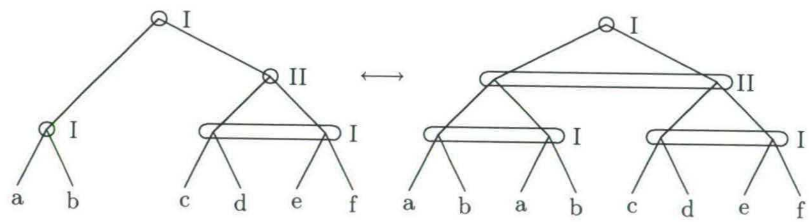


Figure 4.8: Addition of a superfluous move

previous information set that contains predecessors of the information sets defined by the independence condition of $\exists y_{/x}[x = y]$, and in fact already for one single information set there are more than one (see figure 4.7). This is rather unconventional for the games usually studied in game theory, as we remarked in section 4.7: it makes this game one of *imperfect recall*, by a failure of *memory of previous knowledge* (terminology from [Bno04]).

4.8.2 Addition of a superfluous move

A superfluous move may be described as an extra move in the game, the result of which is subsequently forgotten by both players. The addition of such a move to a game, does not alter it, according to this transformation. The transformation is pictured in figure 4.8: a move for player II is added in the left part of the game tree, in such way that it is *superfluous*.

We like to remark that, both in the original picture [Tho52, Fig. 3], and in the version of it in [OR94, Fig. 206.1], the extra move for player II on the left branch is put in the same information set with the move for player II that was already there on the right branch. However, from the formal descriptions of the transformation we do not see where this is required (neither in [Tho52, Definition 16.2, ii], nor in [OR94, p. 207]). This may be due to the fact that both are formulated in the right-left direction of the picture, hence could be more appropriately called 'Removal of a superfluous move'.

From the formal description in [OR94], which is the more detailed one, we understand that the transformation of Addition is not deterministic in the sense that we may either let the superfluous move be a singleton information set, or put it in an appropriate existing information set of the same player (e.g. at least the consistency condition must be satisfied). 'Removal' on the other hand is deterministic (which probably explains why this is the direction formalized).

The following equivalences are instances of this transformation (cf. [vB00b, 12.3.3]):

$$\forall x \exists y_{/x}[x = y] \equiv \forall x \exists z_{/x} \exists y_{/z,x}[x = y]$$

$$\forall x \exists y_{/x}[x = y] \equiv \forall x \exists z \exists y_{/z,x}[x = y]$$

If z does not occur in ψ, we write $\psi_{/z}$ for the result of adding the independence condition $_{/z}$ to all quantifiers and connectives in ψ (this is defined formally in definition 6.2.4 in chapter 6 of this thesis). We may then generalize this to

$$\forall x \exists y_{/x} \psi \equiv \forall x \exists z_{/x} \exists y_{/z,x} \psi_{/z}$$

$$\forall x \exists y_{/x} \psi \equiv \forall x \exists z \exists y_{/z,x} \psi_{/z}$$

Note that the z under the slash of the second existential quantifier is essential, as we know from our recurring example (3.10):

$$\forall x \exists y_{/x}[x = y] \not\equiv \forall x \exists z \exists y_{/x}[x = y]$$

If we would add a superfluous connective move, it would give

$$\forall x \psi \equiv \forall x[\psi_{/\vee} \vee_{/x} \psi_{/\vee}]$$

$$\forall x \psi \equiv \forall x[\psi_{/\vee} \vee \psi_{/\vee}]$$

The latter formulas are not part of our language, because we do not include the possibility of slashing out connectives in the IFG-language, for reasons mentioned in section 4.6. On the other hand, in this situation (like in the formula φ in section 4.6) the connectives under the slash can be interpreted without problem, because both disjuncts $\psi_{/\vee}$ define subtrees of the game tree that have exactly the same structure, player assignment, and outcome function.

It is clear that we cannot simply omit the connectives under the slashes:

$$\forall x \exists y_{/x}[x = y] \not\equiv \forall x[\exists y_{/x}(x = y) \vee \exists y_{/x}(x = y)] \tag{4.3}$$

(Playing in a two-element model, Eloïse does not have a winning strategy for the left formula, while she has one for the right one, as demonstrated in example 4.4.1.)

A correspondence with this transformation may be recognized in the quantifier extraction lemma, which is part of the improved prenex normal form theorem in chapter 6 of this thesis (cf. lemma 6.4.5): if x does not occur in ψ, then

$$(Qx_{/Y}\varphi) \vee \psi \equiv Qx_{/Y}(\varphi \vee \psi_{/x}) \tag{4.4}$$

Here the subgame for ψ becomes preceded by an extra quantifier move $Qx_{/Y}$, but this move is made 'superfluous' by adding $_{/x}$ to all quantifiers and connectives in ψ.

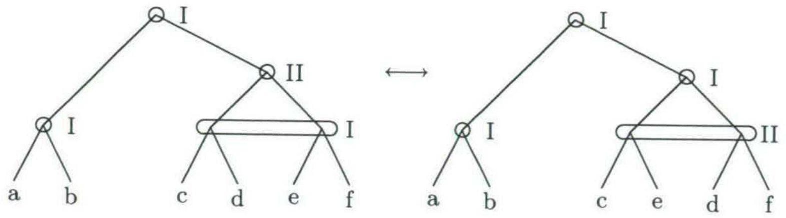


Figure 4.9: Interchange of moves

4.8.3 Interchange of moves

The principle of interchange of moves uses the fact that the order of moves does not matter if the original second player was not informed about the outcome of the first player's move. In figure 4.9 we see how the moves of player II and player I in the right part of the game tree can be interchanged.

This principle is instantiated in the following IFG-example (cf. [vB00b, 12.3.2]):

$$\forall x \exists y_{/x}[x = y] \equiv \exists y \forall x_{/y}[x = y]$$

which can easily be generalized to

$$\forall x \exists y_{/x} \psi \equiv \exists y \forall x_{/y} \psi$$

for arbitrary IFG-formulas ψ.

Note that we also witness an interchange of moves in the quantifier extraction equivalence (4.4) above: in the first two stages leading to the subgame for φ, the quantifier move $(Qx_{/Y})$ is in a sense interchanged with the disjunction move. Interchanging a connective with a quantifier in the opposite direction, like in

$$\forall x[\varphi \vee_{/x} \psi] \equiv (\forall x_{/\vee})\varphi \vee (\forall x_{/\vee})\psi$$

is in principle also possible (it is clear how to interpret the connective under the slashes here), but takes us outside the language again.

4.8.4 Coalescence of moves

The fourth and last transformation, exemplified in figure 4.10, is based on the idea that if a player is assigned two subsequent moves, she will plan ahead and analyze her options for both moves at the same time. In the picture, player I **coalesces** her initial move and her subsequent move in the left part of the tree.

The following example, demonstrating associativity of disjunction, may be seen as an instance of this principle, but also demonstrates a problem with it in our IFG-syntax:

$$(\theta \vee \psi) \vee \chi \equiv \theta \vee (\psi \vee \chi) \equiv \theta \vee \psi \vee \chi$$

Even though the principle seems very straightforward, strictly speaking it takes us outside of our IFG-language! This is unavoidable as in the game trees for our

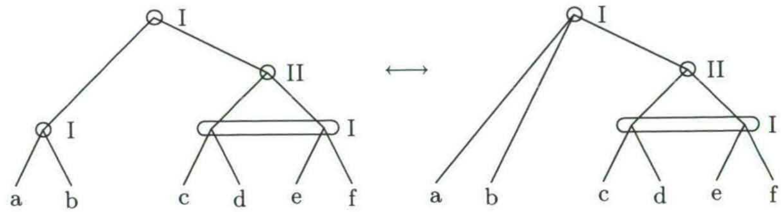


Figure 4.10: Coalescence of moves

semantic IFG-games $\Gamma(\mathfrak{A}, \varphi)$, the number of branches from one node is always either two (in case of a connective), or equal to the cardinality α of $Dom(\mathfrak{A})$. This is dictated by our syntax, which gives rise to only two types of moves: a connective move, or a quantifier move.

Coalescing of moves gives rise to many other possible numbers of branches from one node, like three, in the example above, or $|\mathfrak{A}|^2$ if we would coalesce two quantifier moves. For connectives, the problem may be solved by extending the language with n-ary connectives ($n \in \mathbb{N}, n \geq 2$), so that we can write $\theta \vee \psi \vee \chi$ for example as $\vee(\theta, \psi, \chi)$. For quantifiers, we could allow quantification over tuples of variables. It may be harder to find natural solutions for coalescence of quantifier moves with connective moves. It is not clear to us at this point what kind of logic we would get if we allow such extensions of the language.

4.8.5 Distribution in terms of Thompson transformations

After giving examples of logical equivalence schemes that correspond with (single) Thompson transformations in the previous sections, we now try to see a logical principle (Distribution) in the light of the transformations, in order to find an IFG-version for it.

The propositional distributive law

$$(A \wedge B) \vee C \equiv (A \vee C) \wedge (B \vee C) \tag{4.5}$$

states "that the corresponding games are 'outcome equivalent' in terms of players' 'forcing powers' " ([vB00b, section 1.4]). Indeed: in both the games corresponding to the left and the right formulas respectively, Abélard has the power to let the game end with either an element of the set $\{A, C\}$, or an element of the set $\{B, C\}$, while Eloïse can determine whether the game ends within $\{A, B\}$ or $\{C\}$. [vB00b] brings this up as an example of an "exact 'correspondence' between a logical law and a game-theoretic principle": the distributive law "may be viewed as a transformation allowing us to interchange the order of turns for players in a game without affecting their strategic powers concerning outcomes." In the context of this section on the Thompson transformations, this seems to suggest a correspondence between the (propositional) distributive law and the the Interchange-transformation (section 4.8.3).

However, Theo M.V. Janssen proposed (in a personal correspondence) the following example to demonstrate that for IFG-sentences the distributive law fails.

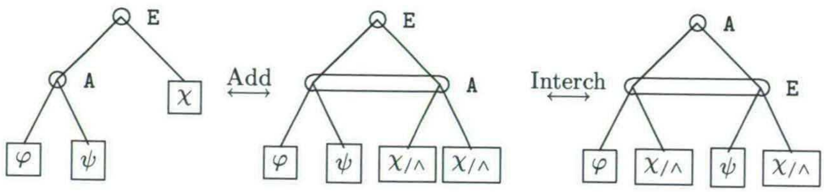


Figure 4.11: Distribution in terms of Addition and Interchange

Example 4.8.1: (failure of straightforward distribution) Let φ be the IFG-sentence

$$\forall x[(E(x) \wedge O(x)) \vee (\exists y_{/x}[x \neq y])]$$

and let $\mathfrak{A}$ be a model with $Dom(\mathfrak{A}) = \mathbb{N}$, the predicate symbol E interpreted by $\{n \in \mathbb{N} | n \text{ is even}\}$, and the predicate symbol O interpreted by $\{n \in \mathbb{N} | n \text{ is odd}\}$. It is not hard to verify that φ is not true in this model: after Abélard's choice of a value for x, Eloïse must choose 'right' at the disjunction (no value for x satisfies the left disjunct $(E(x) \wedge O(x))$). At the right disjunct she has to choose a value for y, distinct from the value of x but also independently of x, which is impossible (there is no constant choice that is distinct from all possible values for x). So, Eloïse has no winning strategy in the game $G_{\mathfrak{A}}(\varphi)$, in other words: φ is not true in $\mathfrak{A}$.

But now consider the formula φ' resulting from applying the distributive law (4.5) to the part of φ under the scope of the universal quantifier:

$$\forall x[(E(x) \vee \exists y_{/x}[x \neq y]) \wedge (O(x) \vee \exists y_{/x}[x \neq y])]$$

Now Eloïse does have a winning strategy in the game $G_{\mathfrak{A}}(\varphi')$: if Abélard's choices put her in the situation of the left conjunct, she can win the game by playing 'left' at the disjunction if the value of x is even, and right otherwise. At the right disjunct $\exists y_{/x}[x \neq y]$, she then knows that the value of x must be odd, so she can choose any even value for y and win. Similarly for the right conjunct.

So, the result φ' of distribution applied to a subformula of φ is true in the model $\mathfrak{A}$, while φ is not true in $\mathfrak{A}$. $\diamond$

This may look puzzling, because the propositional distributive law made us associate distribution with the Thompson transformation of Interchange. But the example above shows how the distributive law for this IFG-formula results in a second occurrence of the subformula $\exists y_{/x}[x \neq y]$, which is associated with Addition. As the non-equivalence (4.3) showed, it is essential that we add the proper independence conditions to moves following an added move (in order to guarantee it is indeed *superfluous*).

If we look at Propositional Distribution in terms of the Thompson transformations, it indeed turns out to be a combination of interchange *and* addition of a superfluous move. This is pictured in figure 4.11 As such, the game tree for the formula $(\varphi \wedge \psi) \vee \chi$ does not allow for the Interchange-transformation, therefore we first transform it into a tree that does, by the Addition-transformation:

$(\varphi \wedge_{/\vee} \psi) \vee (\chi_{/\wedge} \wedge_{/\vee} \chi_{/\wedge})$. In order to get a game to which we can apply Interchange, we put the superfluous move by Abélard (the right conjunction) in one information set with his original move (the left conjunction). If we then apply Interchange, we get the game corresponding to the formula $(\varphi \vee_{/\wedge} \chi_{/\wedge}) \wedge (\psi \vee_{/\wedge} \chi_{/\wedge})$. So:

$$\begin{aligned}(\varphi \wedge \psi) \vee \chi &\equiv (\varphi \wedge_{/\vee} \psi) \vee (\chi_{/\wedge} \wedge_{/\vee} \chi_{/\wedge}) \\ &\equiv (\varphi \vee_{/\wedge} \chi_{/\wedge}) \wedge (\psi \vee_{/\wedge} \chi_{/\wedge})\end{aligned}$$

In the propositional case and if φ, ψ, χ are sentences, the slashed connectives don't seem to be crucial, and the law may be formulated as in the propositional case (4.5): without the connectives under the slashes. (Apparently, the evaluation of propositions and sentences is context independent.) However, the example above demonstrates that this is not the case if we apply distribution at a deeper level in the formula: at the level of subformulas appearing in the scope of previous quantifications, and if those subformulas contain independence conditions regarding those quantifications.

The analysis of distribution based on the Thompson transformations results in the following equivalent for the IFG-sentence φ of example 4.8.1:

$$\forall x[(E(x) \vee_{/\wedge} \exists y_{/x,\wedge}[x \neq y]) \wedge (O(x) \vee_{/\wedge} \exists y_{/x,\wedge}[x \neq y])],$$

rather than in the formula φ' of the example. We conclude that, to express a proper distribution law for IFG-logic, the language would need to contain the possibility to slash for connectives as well as variables.

This is another example showing that we have to be very careful if we want to generalize logical principles from classical logic to IF(G)-logic.

4.8.6 About the status of the transformations

We are aware of the fact that the Thompson transformations are not unproblematic in game theory. They raise questions about the underlying conception of rationality of the players, and the exact interpretation of the notion of information set. In particular, they may turn games of perfect recall into games of imperfect recall, while game theory usually restricts its scope to the first type of games only. A detailed discussion about problems with transformations under different conceptions of information set, can be found in [ER94, section 2]. Also, [OR94, p. 209] mentions that the conception of rationality underlying the transformations ignores so-called "framing effects", while psychologists have found "that even minor variations in the framing of a problem may dramatically affect the participants' behavior (see for example Tcersky and Kahneman (1986))."

In [SP01, p. 211], Sandu and Pietarinen argue on such grounds that the Thompson transformations "cannot be taken to form a basis for a complete calculus for IF [propositional] logic (as suggested by [vB00a])". This conclusion may be true, but we are not sure about the arguments. On the one hand, we conjecture that the noted *strategic essentiality* demonstrated in [ER94], does not occur in win-loss games, in which no probabilities are involved. On the other hand, *imperfect recall* is already omnipresent in semantic games for IF(G)-logic, as we will saw in section 4.7.

Without the expectations to find a (complete) calculus for IF-logic, we found some clear and interesting correspondences with the transformations in our logic. The fact that some of these corresponding logical facts are also a bit puzzling, could even be explained through the game-theoretic discussions about the transformations.

Throughout this section, we encountered several situations witnessing that the the class of semantic games for our IFG-language is not closed under the Thompson transformations. There were two main causes for this: on the one hand, because we excluded independence of connective moves, as this could in general lead to problems of interpretation (see section 4.6). On the other hand, because the game trees corresponding with the semantic games for IFG-formulas have only two types of branching: in two or in the cardinality α of the domain of the model.

The latter disqualified the transformation of *coalescence* in IF-logic. This transformation is essential for Thompson's result that we can transform any extensive form game into reduced normal form by application of the right transformations in the right order. Coalescence must be essential, because it is the only transformation reducing the depth of the game tree (and a reduced normal form for a two-player game has depth 2).

This shows that even though we found some nice correspondences between logical equivalence schemes and the Thompson transformations, the latter can not deliver a syntactic 'reduced normal form' for IF-sentences. In that sense, the transformations do not contribute to IF-logic what made them relevant for the theory of games in extensive form in general. On the other hand, it does not seem to be too difficult to be more liberal in the translation of an IFG-formula into a game tree, e.g. by allowing a sequence of quantifiers of the same polarity to count as one move. One could even wonder what logic we get if we used the Thompson transformations as 'axiom schemes', and designed a language that did suit them. (This question was asked by Wilfrid Hodges at the Workshop on Knowledge and Games in Liverpool, July 10-11, 2004. We do not have any specific directions towards an answer yet. More generally, in in [vB01], van Benthem argues that a dynamic epistemic language would be natural to model extensive games, at least if analyzed on the 'action level'.)

Conversely, given a logical normal form (prenex normal form, Skolem normal form, Distributive/Conjunctive normal form), it would be interesting to see which game-theoretic transformations correspond to the syntactic transformations used to rewrite a formula into that normal form.

4.9 Conclusions

We introduced a generalization of the IF-language, the language $\mathcal{L}_{IFG}$. We feel this language corresponds naturally to the semantic games as we understand them from their rules: there are two types of moves (quantifier and connective), negation determines which of the two players makes the moves.

The modeling of semantic games in game-theoretic terms, makes it possible to characterize them as finite depth, two-player, win-loss games of imperfect information, and possibly, of imperfect recall. Strategies can be seen as sets of choice functions, that correspond to the 'inside-out' Skolem functions of our previous chapter.

The game trees (as constituted by the sets of histories, that are in turn determined by the syntax of the formula and the domain of the model) and the information partitions (as determined by the slash operator) have very specific structures. However, it seems hard to give a game-theoretic characterization of these structures. We would get a more general class of game structures and information partitions by introducing independence of connectives. But we showed that in most cases, the interpretation of connectives under the slash operator does not add new imperfect information. For some cases where it does, we would have to revise the treatment of strategies as sets of choice functions.

We have pointed out that, while game theory usually presupposes the opposite, many semantic games for IF(G)-sentences are of *imperfect recall*, not only in the sense that players may forget their own previous *actions*, but also in the sense that players may forget what they *knew* before. Hodges' example gives rise to games of imperfect recall in the latter sense.

In the last section, we have given some equivalence schemes for IFG-sentences that can be seen to correspond with the Thompson transformations from game theory. We have not tried to be complete, so we expect there to be more. On the other hand, we also noticed that the transformations also take us beyond what is syntactically expressible in the IFG-language. In that sense, the application of the transformations cannot be expected to bring us a normal form for IF-formulas that corresponds with the reduced normal form of extensive games.

The observations in this chapter make us believe that semantic games are not a natural subclass of all games in extensive form, and therefore the use of game-theoretic results for IFG-logic is limited. But we consider the formalization of the semantic games for IFG-formulas into the existing game-theoretic terminology as a useful result in itself. Other than the definition of semantic games in terms of their rules, it yields mathematical objects to interpret central concepts as *imperfect information* and a strategy.

Chapter 5

Satisfaction for open IFG-formulas

In this chapter, we present and comment on a semantics for IFG-formulas with free variables, based on the semantics of [CK99]. We first define the game semantics, and then a set of inductive clauses that define the same notion of satisfaction. We show how the original existential clause needed to be improved, and prove that the (corrected) semantics makes IFG-logic a conservative extension of first order logic (for arbitrary formulas). Finally, we demonstrate that IF-logic (with the slashing convention) fails to be a conservative extension of first order logic (sentences only), if we do not require the sentences to be *regular*. This serves as preparation for the type of problems we found in the prenex normal form theorem of [CK99], which we discuss and solve in the next chapter.

The contents of both the current and the next chapter are based on joint work with Theo M.V. Janssen and Xavier Caicedo ([JD05],[CDJ]).

5.1 Independence with free variables

While the language of IF-logic consists of sentences only, the IFG-language also contains formulas with free variables. Nevertheless, in the previous chapter, we only formalized semantic games for IFG-*sentences* (cf. section 4.4). Open IFG-formulas only occurred as a subformula ψ of a given sentence φ. This sentence provided a context for the subformulas, that determined the player distribution (through the polarity of ψ), and the available information (the relative domain X_ψ^φ determined which variables have a value assigned when we reach ψ in a game for φ).

For first order logic, it is no problem to define game theoretical semantics for a formula ψ with free variables: just add a valuation v for the free variables as extra parameter to the semantic game, and let the game start from the initial position $\langle \psi, v, 1\rangle$ (cf. definition 2.4.1). This initial position defines a subtree of the game for any sentence containing ψ as a subformula. Because first order formulas define games of perfect information, there are no information sets involved. If a player can win from this initial position, she can win from this position in the context of any game tree. Hence, this subtree can be regarded apart from its context without changing 'its meaning'.

But the situation is different if we want to preserve the 'meaning' of the slash operator. The meaning of the slash operator is not determined by a single position in the game, it is determined by the subdivision of a certain *set of positions* into information sets, as induced by the slash operator. To illustrate this, consider the IFG-formula $\exists y_{/x}[x = y]$, with x as its only free variable occurring under the slash. Suppose we would play a semantic game for this formula, with only the valuation $(x\colon 0)$ as extra parameter. If Eloïse would choose the value 0 for y, there would be no way of telling from the resulting valuation $(xy\colon 00)$ that y was chosen independently of x. But if we take a *set of valuations* to begin with, the set $\{(x\colon 0), (x\colon 1)\}$ for example, the situation becomes different: if Eloïse chooses the value 0 from the position with $(x\colon 0)$, and 1 from the position with $(x\colon 1)$, her choice obviously depends on the value assigned to x. However, if she would choose the same value in both cases, her choice would be *independent* of the value assigned to x.

The idea to use a set of valuations in order to interpret open formulas with the independence operator, was introduced by Wilfrid Hodges' trump semantics, in [Hod97a]. Trump semantics coincides with game theoretical semantics for IFG-sentences. Hence, by presenting his semantics, Hodges challenged Hintikka's claim that logics with the slash operator are non-compositional by nature. As mentioned in our introductory chapter 1, the presentation of trump semantics led to interesting discussions about degrees of compositionality ([SH01],[Hod01b]), but we do not go into that matter here.

The semantics we present and comment upon in this chapter is a variation of Hodges' trump semantics, as formulated by Xavier Caicedo and Michal Krynicki in [CK99]. Their version came into existence in an attempt to get a better understanding of trump semantics. Their paper also demonstrates why (in some sense) compositional semantics for open formulas are useful. Until now, we have encountered some equivalence schemes that hold for logics with the slash operator: the ones associated with negation (cancellation of double negation, De Morgan's laws), and some schemes associated with the Thompson transformations (but with all the remarks about them, one may distrust the transformations as formal justification). With a decent notion of equivalence for open formulas, and some kind of substitution principle for equivalent subformulas, it becomes easier to formulate *and* prove general equivalence schemes, without having to rely on translation procedures to other systems (game theory, or Σ^1_1). This is demonstrated by the prenex normal form theorem given in [CK99] (with some interesting flaws however, which will be the topic of our next chapter).

In this chapter, we will introduce the game semantics for open IFG-formulas as defined in [CK99], and the (corrected) clauses for an inductive notion of satisfaction equivalent to this game semantics. As a first exercise with this semantics, we prove that IFG-logic with this semantics for open formulas is (in a specific sense) a conservative extension of first order logic.

Continuing the theme of conservative extension, and preparing for the type of situations we will encounter in the next chapter, we pay some attention to signaling phenomena. We do so in relation to the slashing convention of IF-logic, and show that nested quantification over the same variable may cause first order sentences to become *undetermined* by the slashing convention. Thereby, IF-logic is only a conservative extension of what we will call *regular* first order sentences.

5.2 Game semantics for open formulas

The language $\mathcal{L}_{IFG}$ used in [CK99] was already defined in section 4.3. We now give the definition of game semantics for IFG-formulas with free variables. The definition has to account for independence of variables whose assignment was determined outside the game, as in $\mathcal{L}_{IFG}$ variables may occur free under the slashes in the formula. The ingredient we have to add to the game in order to be able to interpret this independence, is a *set* of valuations (rather than one single valuation for $Fv(\varphi)$, as we could do for first order formulas). For the rest, the games for (open) IFG-formulas are very similar to the semantic games as defined in definition 2.4.1. The main aim of this section is to introduce the games in the terminology of [CK99, p. 18–19], because we will adopt this terminology to comment on the results of that paper.

We first introduce some useful notations concerning sets of valuations. Recall that we use the letter A to denote the domain of a given model $\mathfrak{A}$.

Notation 5.2.1 (sets of variants) *(Cf. definition 2.3.4.) Let $\mathfrak{A}$ be a model, X a set of variables, and $V \subseteq A^X$. If $a \in A$, then $V_{x:\,a} := \{v(x\colon a) | v \in V\}$. Also, we define $V_{x:\,A} := \{v(x\colon a) | v \in V, a \in A\}$, i.e. the set of all x-variants (in A) of the valuations in $V \subseteq A^X$.*

We generalize this notation to sets of variables: for a finite set of variables Y, we write $V_{Y:\,A} \subseteq A^{X \cup Y}$ for the set of all 'Y-variants' of the elements of V. An Y-variant of $v \in V \subseteq A^X$ is a valuation $w \in A^{X \cup Y}$ such that $v(x) = w(x)$ for all $x \in X \setminus Y$. We will only use this notation once (in definition 5.2.4)

Notation 5.2.2 (explicit notation for sets of valuations) *(Extending notation 2.3.2.) If $V \subseteq A^{\{x_1,\dots,x_k\}}$ is a finite set of valuations, say $V = \{v_1, \dots, v_n\}$ with $v_j = (x_1 \dots x_k\colon a_{1j} \dots a_{kj})$, we write V as*

$$\{x_1 \dots x_k\colon (a_{11} \dots a_{k1}), \cdots, (a_{1n} \dots a_{kn})\}.$$

For example, we write $\{0,1\}^{\{x,y\}}$ as $\{xy\colon 00, 01, 10, 11\}$.

Let us now define the semantic games for open IFG-formulas. The main ingredient of a game is an IFG-formula. The aim of the game is to determine whether a formula φ is satisfied in a given suitable model $\mathfrak{A}$, with respect to a given set of suitable valuations $V \subseteq A^X$ (i.e. with $Fv(\varphi) \subseteq X$). As before, Eloïse tries to confirm satisfaction, while Abélard tries to refute it. The existence of a winning strategy for one of the players, determines whether or not the formula is satisfied in $\mathfrak{A}$ with respect to V.

We first define how the game is played: which player has to move in a given position, what are his/her possible moves, and what is the effect of the move. Following the syntax of the initial IFG-formula φ, the players will encounter subformulas $\psi \vee_{/Y} \theta$ or $\exists x_{/Y}\psi$. The subscript indicates that the choice associated with that quantifier or connective has to be made independently of the variables in Y. This is a restriction on the motivation for the choice, but not on the available choices. Therefore in the description of the plays, it makes no difference whether $_{/Y}$ occurs as subscript or not. Its role will be defined in the definition of the notion of the *choice functions* that constitute the strategies (cf. definition 5.2.7).

To describe the *positions* for a subformula ψ of φ in the game $\mathtt{G}(\mathfrak{A}, \varphi, V)$, we need to determine the domain of the valuations for ψ. This is determined by the set of variables in the scope of whose quantification ψ occurs in φ. Recall from definition 4.3.5 that we denote this set of variables as X^{φ}_{ψ}, which we can read as the set of variables to which the players assign (new) values in a play of $\mathtt{G}(\mathfrak{A}, \varphi, V)$, before they 'reach' the subformula ψ. Because every play starts with a valuation with domain X, this makes $X \cup X^{\varphi}_{\psi}$ the domain for the valuations in the positions for ψ.

Compared with the positions in definition 2.4.1 of semantic games for first order sentences, we replace the parameter p by the notion of polarity, as defined in definition 4.3.7. We use the following notation:

Notation 5.2.3 (positive and negative subformulas) *Let φ be an IFG-formula. Recall that $S(\varphi)$ is the set of subformulas of φ without initial negation signs (cf. notation 4.3.3). Let $S^+(\varphi) := \{\psi \in S(\varphi) | \psi$ is positive in $\varphi\}$ and $S^-(\varphi) := \{\psi \in S(\varphi) | \psi$ is negative in $\varphi\}$.*

Definition 5.2.4 (IFG-semantic games) *Given an IFG-formula φ, a suitable model $\mathfrak{A}$, and a set of valuations $V \subseteq A^X$ in $\mathfrak{A}$ with $Fv(\varphi) \subseteq X$, we define the game $\mathtt{G}(\mathfrak{A}, \varphi, V)$ as follows:*

A **position** *in the game is a pair $\langle \psi, v \rangle$ where $\psi \in S(\varphi)$, and $v \in V_{X^{\varphi}_{\psi}:A} \subseteq A^{X \cup X^{\varphi}_{\psi}}$. A position $\langle \psi, v \rangle$ is a* **terminal** *if ψ is atomic.*

The moves (by the players Abélard and Eloïse) determine the transitions from a non-terminal position $\langle \psi, v \rangle$ to the next position. The polarity of ψ in φ determines the **player assignment***: if ψ is positive in φ ($\psi \in S^+(\varphi)$), then the position is assigned to Eloïse; if ψ is negative in φ ($\psi \in S^-(\varphi)$), then it is assigned to Abélard.*

- *if $\psi \equiv \exists x_{/Y}\psi'$, then the assigned player chooses an $a \in A$, and the play continues from position $\langle |\psi'|, v(x\colon a) \rangle$.*
- *if $\psi \equiv \psi_l \vee_{/Y} \psi_r$, then the assigned player chooses a subformula ψ' from $\{\psi_l, \psi_r\}$, and the play continues from position $\langle |\psi'|, v \rangle$*

A **play** *of the game is a terminal root path in the game tree, i.e. a sequence of positions that results from subsequent moves, starting with a position $\langle \varphi, v \rangle$ for some $v \in V$, and ending with a terminal position.*

The terminal position $\langle\psi, v'\rangle$ of a play, determines the **outcome** *of the play: the play is winning for Eloïse in the following cases:*

- $\psi \in S^+(\varphi)$ *and* $\mathfrak{A} \models \psi[v']$, *or*
- $\psi \in S^-(\varphi)$ *and* $\mathfrak{A} \not\models \psi[v']$.

In the other two cases, Abélard wins the play.

The **semantic game** $\mathtt{G}(\mathfrak{A}, \varphi, V)$ *is the collection of all plays starting from a position $\langle\varphi, v\rangle$ with $v \in V$, together with the player assignment to all non-terminal positions and the outcome function for all terminal positions.*

Now we know how the game is played, we can define the central notion of (winning) strategy. We do so in terms of *choice functions.* A choice function f_ψ is a function that describes which choice the assigned player makes at ψ in any of the positions $\langle\psi, v\rangle$ (where $v \in V_{X^\varphi_\psi : A}$). The requirement that a choice may not depend on the variables in a set Y appearing under the slash, is formalized by requiring that f_ψ yields the same choice for two valuations that 'coincide out of Y':

Definition 5.2.5 (to coincide out of Y) *Let $v, w \in A^X$ be two valuations of a set of variables X in a model $\mathfrak{A}$, and Y an arbitrary set of variables. We then say that v, w* **coincide out of** *Y ($v \sim_Y w$) if and only if $v(x) = w(x)$ for all $x \in X \setminus Y$. (Note that $\sim_Y$ defines an equivalence relation on V.)*

Definition 5.2.6 (independence of Y) *If $V \subseteq A^X$ and f is a function with domain V, then f* **is independent of** *Y (or f is Y-independent) (on its domain V) if for all $v, w \in V$: $v \sim_Y w$ implies $f(v) = f(w)$.*

Note that any function is independent of the empty set of variables.

Definition 5.2.7 (choice functions) *Let $\psi \in S(\varphi)$ be a subformula of φ. A* **choice function** *f_ψ in the game $\mathtt{G}(\mathfrak{A}, \varphi, V)$ is a function s.t.*

- *if $\psi \equiv \exists x_{/Y}\psi'$, then $f_\psi : A^{X \cup X^\varphi_\psi} \to A$ is Y-independent; or*
- *if $\psi \equiv \psi_l \vee_{/Y} \psi_r$, then $f_\psi : A^{X \cup X^\varphi_\psi} \to \{l, r\}$ is Y-independent*

Definition 5.2.8 (strategies) *A strategy for Eloïse in the game $\mathtt{G}(\mathfrak{A}, \varphi, V)$ is a set of choice functions $\{f_\psi\}_{\psi \in S^+(\varphi)}$, one for each positive non-atomic subformula ψ. Similarly, a strategy for Abélard is a set of choice functions f_ψ, one for each negative non-atomic subformula $\psi \in S^-(\varphi)$.*

A strategy is called a **winning strategy** *in game $\mathtt{G}(\mathfrak{A}, \varphi, V)$ if playing in accordance with that strategy guarantees the owner of the strategy to win the game from all starting positions $\langle\varphi, v\rangle$ ($v \in V$).*

Definition 5.2.9 (game satisfaction) *Let φ be an IFG-formula, $\mathfrak{A}$ a suitable model, and $V \subseteq A^X$ a set of suitable valuations for φ in $\mathfrak{A}$ (i.e. with $Fv(\varphi) \subseteq X$). We then define positive and negative game satisfaction as follows:*

$\mathfrak{A} \models^+_{\mathtt{G}} \varphi[V]$ *('φ is true in $\mathfrak{A}$ with respect to V')*
iff there exists a winning strategy for Eloïse in $\mathtt{G}(\mathfrak{A}, \varphi, V)$
$\mathfrak{A} \models^-_{\mathtt{G}} \varphi[V]$ *('φ is false in $\mathfrak{A}$ with respect to V')*
iff there exists a winning strategy for Abélard in $\mathtt{G}(\mathfrak{A}, \varphi, V)$.

5.3 Inductive clauses for satisfaction

The game theoretic definition for positive and negative satisfaction (in other words: truth and falsity) of IFG-formulas is one in terms of the existence of a set of functions that satisfy independence conditions determined by the slash operator, as we saw in the previous section. Hence, if we want to show that a given formula is true or false with respect to a certain model $\mathfrak{A}$ and set of valuations V, we have to come up with a set of functions that witnesses this.

However, the definition given above does not yet give us a mathematically precise account of what it means that some strategy is applied by a player, and how to verify that it is winning. This makes it hard to formally prove statements on the evaluation of formulas, particularly if they claim that no winning strategies exist.

In this section we give a set of clauses that enables us to read game satisfaction for IFG as an inductive notion of satisfaction. We first present them as an alternative satisfaction definition $\models^+_{\mathtt{I}}$ and then prove that the notions of $\models^+_{\mathtt{I}}$ and $\models^+_{\mathtt{G}}$ coincide.

The inductive clauses are a useful tool in proving equivalence schemes for IFG-logic. In this form, our semantics resembles the classical Tarskian definition of satisfiability for first order logic, except that we have to work with *sets* of valuations rather than single valuations. The clauses given here are an improved version of the clauses given in lemma 1.1 of [CK99, p. 21].

We will use the notion of Y-saturatedness for sets of valuations (from [CK99]) as the counterpart of the notion of independence for choice functions.

Definition 5.3.1 (Y-saturatedness) *If $V \subseteq A^X$ is a set of valuations, and $V' \subseteq V$ a subset of V that is closed under the equivalence relation $\sim_Y$ on V (i.e. if $v \in V'$ then for any $w \in V$ with $v \sim_Y w$: $w \in V'$), we say that V'* **is Y-saturated in** *V. Note that the Y-saturated subsets of $V \subseteq A^X$ are unions of equivalence classes of the relation $\sim_Y$ on V.*

It is easy to verify that if V_1 and V_2 are Y-saturated subsets of V, then $V_1 \cup V_2$, $V_1 \cap V_2$, and $V_1 \setminus V_2$ are also Y-saturated subsets of V.

Definition 5.3.2 *The partition $V/_{\sim_Y}$ of V into its $\sim_Y$ equivalence classes is called the* **Y-saturated partition** *of V.*

We are now prepared to define an inductive notion of satisfaction for IFG, based on Lemma 1.1 in [CK99, p. 21]:

Definition 5.3.3 (inductive satisfaction) *Let φ be an IFG-formula, $\mathfrak{A}$ a suitable model with domain A, and $V \subseteq A^X$ a set of suitable valuations (i.e. $Fv(\varphi) \subseteq X$). Then $\mathfrak{A} \models^{\pm}_{\mathrm{I}} \varphi[V]$ is defined inductively as follows:*

(At) *If φ is atomic, then*
$\mathfrak{A} \models^{+}_{\mathrm{I}} \varphi[V]$ *iff* $\mathfrak{A} \models \varphi[v]$ *for all* $v \in V$.
$\mathfrak{A} \models^{-}_{\mathrm{I}} \varphi[V]$ *iff* $\mathfrak{A} \not\models \varphi[v]$ *for all* $v \in V$.

($\neg$) $\mathfrak{A} \models^{+}_{\mathrm{I}} \neg\psi[V]$ *iff* $\mathfrak{A} \models^{-}_{\mathrm{I}} \psi[V]$
$\mathfrak{A} \models^{-}_{\mathrm{I}} \neg\psi[V]$ *iff* $\mathfrak{A} \models^{+}_{\mathrm{I}} \psi[V]$

($\vee, +$) $\mathfrak{A} \models^{+}_{\mathrm{I}} (\psi \vee_{/Y} \theta)[V]$ *iff* $\mathfrak{A} \models^{+}_{\mathrm{I}} \psi[V_1]$ *and* $\mathfrak{A} \models^{+}_{\mathrm{I}} \theta[V_2]$ *for some* $V_1, V_2 \subseteq V$, *both Y-saturated in V, with* $V = V_1 \cup V_2$.

($\vee, -$) $\mathfrak{A} \models^{-}_{\mathrm{I}} (\psi \vee_{/Y} \theta)[V]$ *iff* $\mathfrak{A} \models^{-}_{\mathrm{I}} \psi[V]$ *and* $\mathfrak{A} \models^{-}_{\mathrm{I}} \theta[V]$.

($\exists, +$) $\mathfrak{A} \models^{+}_{\mathrm{I}} \exists x_{/Y}\psi[V]$ *if and only if there is a partition* $V = \cup_{i \in I} V_i$ *into Y-saturated subsets V_i of V, and there is for each $i \in I$ an $a_i \in A$, such that* $\mathfrak{A} \models^{+}_{\mathrm{I}} \psi[\cup_{i\in I}(V_i)_{x:\, a_i}]$

($\exists, -$) $\mathfrak{A} \models^{-}_{\mathrm{I}} \exists x_{/Y}\psi[V]$ *iff* $\mathfrak{A} \models^{-}_{\mathrm{I}} \psi[V_{x:\, A}]$

Before we prove this notion of satisfaction to coincide with game satisfaction, we make some remarks:

1. Note that the right hand side of the atomic clauses uses the classical (Tarski) definition of satisfaction.

2. Note that if we let $V = \emptyset \subseteq A^X$ be the empty set of valuations, this inductive definition of satisfaction yields:

$$\mathfrak{A} \models^{\pm}_{\mathrm{I}} \varphi[\emptyset], \text{ for any IFG-formula } \varphi$$

   This might look anomalous, but it is actually necessary for the situation with disjunction, where the empty sets of valuations can occur if V is split into V and $\emptyset$. (We remark that Hodges' trump semantics shares this feature, cf. [Hod97c, p. 57].) Be aware that this is different from saying that formulas are always satisfied by the singleton set $A^{\emptyset} = \{\lambda\}$: this is not the case. In fact, satisfaction with respect to $A^{\emptyset}$ is only *defined* for formulas with no free variables, i.e. sentences.

3. Compare the requirements for the positive clauses for disjunction and existential quantification. In the case of disjunction, we do not require the V_i to be disjoint, while a choice *function* $f : V \to \{l, r\}$ would divide V into two disjoint sets: $f^{-1}(\{l\})$ and $f^{-1}(\{r\})$. It follows from the next lemma, that requiring disjointness here does not make a difference.

   In the case of quantification, we do require disjointness as we *partition* V into Y-saturated subsets. Note that if we take the Y-saturated partition of V, this can be seen to correspond to the information partition of the nodes in the extensive game tree that are associated with $\exists x_{/Y}\psi$ (if we see $\exists x_{/Y}\psi$

as subformula of some sentence φ, and if V is the set of all valuations at the nodes associated with $\exists x_{/Y}\psi$). For each i, a_i is then the unique action prescribed by a strategy for all elements in the information set V_i. We come back to the existential clause in section 5.4.

We now prove that satisfaction as defined in definition 5.3.3 coincides with game satisfaction (as in [JD05]). The proof may also explain the intuitions behind the inductive clauses in terms of the games. In this respect, the following definition (extending notation 5.2.1) is useful:

Definition 5.3.4 (x-variations) *Let $V \subseteq A^X$ be a set of valuations and x a variable. A subset V_x of $V_{x:\,A}$ is called* **an x-variation of V** *if V_x contains* at least one *x-variant for every $v \in V$. If $f : V \to A$, we define* **the x-variation of V by f** *as follows:*

$$V_{x:\,f} := \{v(x\colon f(v))|v \in V\}.$$

A set $V_{x:\,f}$ can be seen as the set of all possible valuations the game continues with if Eloïse applies the choice function f to choose x in the game $\mathtt{G}(\mathfrak{A}, \exists x_{/Y}\varphi, V)$. Notice that x may or may not be in $X = dom(V)$.

Theorem 5.3.5 (game- and inductive satisfaction coincide) *Suppose φ is an IFG-formula and $\mathfrak{A}$ a suitable model. Then for any set $V \subseteq A^X$ of valuations for $\varphi \in \mathfrak{A}$:*

$$\mathfrak{A} \models^{\pm}_{\mathtt{G}} \varphi[V] \text{ iff } \mathfrak{A} \models^{\pm}_{\mathtt{I}} \varphi[V]$$

Proof: By induction on the syntax of φ:

(At) If φ is atomic, the semantic game $\mathtt{G}(\mathfrak{A}, \varphi, V)$ is the collection of one-position plays consisting only of a terminal position $\langle \varphi, v\rangle$ with $v \in V$. Because strategies in such a game are empty for both players, it follows that Eloïse has a winning strategy ($\mathfrak{A} \models^{+}_{\mathtt{G}} \varphi[V]$) if and only if $\mathfrak{A} \models \varphi[v]$ for all $v \in V$ (see the definition of the outcome of a play in definition 5.2.4), i.e. $\mathfrak{A} \models^{+}_{\mathtt{I}} \varphi[V]$. Similarly, $\mathfrak{A} \models^{-}_{\mathtt{G}} \varphi[V]$ if and only if $\mathfrak{A} \not\models \varphi[v]$ for all $v \in V$, i.e. $\mathfrak{A} \models^{-}_{\mathtt{I}} \varphi[V]$.

($\neg$) Suppose $\varphi = \neg\psi$. $\mathtt{G}(\mathfrak{A}, \neg\psi, V)$ is the same game as $\mathtt{G}(\mathfrak{A}, \psi, V)$, except that the polarity of the subformulas is reversed: $S^{+}(\neg\psi) = S^{-}(\psi)$ and $S^{-}(\neg\psi) = S^{+}(\psi)$. It follows that any winning strategy for Eloïse in $\mathtt{G}(\mathfrak{A}, \neg\psi, V)$ is a winning strategy for Abélard in $\mathtt{G}(\mathfrak{A}, \psi, V)$ and conversely. Hence:

$$\begin{array}{lll} \mathfrak{A} \models^{+}_{\mathtt{G}} \neg\psi[V] & \text{iff} & \mathfrak{A} \models^{-}_{\mathtt{G}} \psi[V] \\ & \text{iff (by induction)} & \mathfrak{A} \models^{-}_{\mathtt{I}} \psi[V] \\ & \text{iff (by definition)} & \mathfrak{A} \models^{+}_{\mathtt{I}} \neg\psi[V] \end{array}$$

Similarly, $\mathfrak{A} \models^{-}_{\mathtt{G}} \neg\psi[V]$ iff $\mathfrak{A} \models^{-}_{\mathtt{I}} \neg\psi[V]$.

($\vee$) Suppose $\varphi = \psi \vee_{/Y} \theta$.

(+) First, suppose $\mathfrak{A} \models^+_{\mathtt{G}} \psi \vee_{/Y} \theta[V]$ and let $f = f_\varphi$ be the 'first' choice function of a winning strategy F_φ for Eloïse in $\mathtt{G}(\mathfrak{A}, \varphi, V)$. Now define $V_1 = f^{-1}(\{l\})$ and $V_2 = f^{-1}(\{r\})$. Then $\mathfrak{A} \models^+_{\mathtt{G}} \psi[V_1]$, $\mathfrak{A} \models^+_{\mathtt{G}} \theta[V_2]$, and $V = V_1 \cup V_2$. It is easy to verify that because f is independent of Y, both V_i are Y-saturated in V: if $v \sim_Y w$, then $f(v) = f(w)$, hence v, w are in the same V_i. By induction hypothesis we have $\mathfrak{A} \models^+_{\mathtt{I}} \psi[V_1]$ and $\mathfrak{A} \models^+_{\mathtt{I}} \theta[V_2]$, hence $\mathfrak{A} \models^+_{\mathtt{I}} \varphi[V]$.
Conversely, suppose $\mathfrak{A} \models^+_{\mathtt{I}} \varphi[V]$ and determine V_1, V_2, Y-saturated in V with $V_1 \cup V_2 = V$ such that $\mathfrak{A} \models^+_{\mathtt{I}} \psi[V_1]$ and $\mathfrak{A} \models^+_{\mathtt{I}} \theta[V_2]$. By induction hypothesis we have winning strategies F_ψ for Eloïse in $\mathtt{G}(\mathfrak{A}, \psi, V_1)$ and F_θ for $\mathtt{G}(\mathfrak{A}, \theta, V_2)$. Now define $f_\varphi : V \to \{l, r\}$ by $f_\varphi(v) = l$ if $v \in V_1$ and $f_\varphi(v) = r$ otherwise (i.e. if $v \in V \setminus V_1$). Then $F_\varphi = F_\psi \cup F_\theta \cup \{f_\varphi\}$ is a winning strategy for Eloïse in the game $\mathtt{G}(\mathfrak{A}, \varphi, V)$, hence $\mathfrak{A} \models^+_{\mathtt{G}} \varphi[V]$.

(-) The following are equivalent:

* $\mathfrak{A} \models^-_{\mathtt{G}} \psi \vee_{/Y} \theta[V]$.
* $\mathfrak{A} \models^-_{\mathtt{G}} \psi[V]$ and $\mathfrak{A} \models^-_{\mathtt{G}} \theta[V]$ (a winning strategy for Abélard in $\mathtt{G}(\mathfrak{A}, \varphi, V)$ does not contain a choice function for φ, so it must be a winning strategy in both $\mathtt{G}(\mathfrak{A}, \psi, V)$ and $\mathtt{G}(\mathfrak{A}, \theta, V)$).
* $\mathfrak{A} \models^-_{\mathtt{I}} \psi[V]$ and $\mathfrak{A} \models^-_{\mathtt{I}} \theta[V]$ (by induction hypothesis)
* $\mathfrak{A} \models^-_{\mathtt{I}} \psi \vee_{/Y} \theta[V]$ (by definition)

($\exists$) Suppose $\varphi = \exists x_{/Y} \psi$.

(+) First, suppose $\mathfrak{A} \models^+_{\mathtt{G}} \exists x_{/Y} \psi[V]$ and let $f = f_\varphi$ be the 'first' choice function of a winning strategy F_φ for Eloïse in $\mathtt{G}(\mathfrak{A}, \varphi, V)$. Let $V_{x:\,f}$ be the x-variation of V by f, as in definition 5.3.4. Then $\mathfrak{A} \models^+_{\mathtt{G}} \psi[V_{x:\,f}]$ by the strategy $F_\psi = F_\varphi \setminus \{f\}$. By induction we then also have $\mathfrak{A} \models^+_{\mathtt{I}} \psi[V_{x:\,f}]$.
Now let $V = \cup_i V_i$ be the Y-saturated partition of V, then, because f is independent of Y, $f(v) = f(w)$ for all v, w in the same V_i. So, we can define a_i to be $f(v)$ for an arbitrary $v \in V_i$, and then: $\cup_i (V_i)_{x:\,a_i} = V_{x:\,f}$. Because $\mathfrak{A} \models^+_{\mathtt{I}} \psi[V_{x:\,f}]$ we may now conclude $\mathfrak{A} \models^+_{\mathtt{I}} \exists x_{/Y} \psi[V]$.
Conversely, assume $\mathfrak{A} \models^+_{\mathtt{I}} \exists x_{/Y} \psi[V]$, and determine a Y-saturated partition $\cup_i V_i$ of V and $a_i \in A$ such that $\mathfrak{A} \models^+_{\mathtt{I}} \psi[\cup_i (V_i)_{x:\,a_i}]$. By induction: $\mathfrak{A} \models^+_{\mathtt{G}} \psi[\cup_i (V_i)_{x:\,a_i}]$, so determine a winning strategy F_ψ for Eloïse in $\mathtt{G}(\mathfrak{A}, \psi, \cup_i (V_i)_{x:\,a_i})$. Now define $f : V \to A$ such that for each $v \in V_i$ $f(v) = a_i$. The fact that the V_i form a partition, makes f well-defined on V, and because the V_i are Y-saturated, f is automatically Y-independent. Furthermore, $V_{x:\,f} = \cup_i (V_i)_{x:\,a_i}$. Thus, $F_\varphi = F_\psi \cup \{f\}$ is a winning strategy for Eloïse in the game $\mathtt{G}(\mathfrak{A}, \varphi, V)$, hence $\mathfrak{A} \models^+_{\mathtt{G}} \varphi[V]$.

(-) The following are equivalent:

* $\mathfrak{A} \models^-_{\mathtt{G}} \exists x_{/Y} \psi[V]$.
* $\mathfrak{A} \models^-_{\mathtt{G}} \psi[V_{x:\,A}]$ (a winning strategy for Abélard in $\mathtt{G}(\mathfrak{A}, \varphi, V)$ does not contain a choice function for φ, it must be a winning strategy for all possible assignments to x chosen by Eloïse)

* $\mathfrak{A} \models^-_{\mathrm{I}} \psi[V_{x:\,A}]$ (by induction hypothesis)
* $\mathfrak{A} \models^+_{\mathrm{I}} \exists x_{/Y}\psi[V]$ (by definition)

◁

In a sense, the definition of $\models^+_{\mathrm{I}}$ makes the effect of the application of one choice function in a (winning) strategy explicit in terms of sets of valuations. This will be very useful when we want to prove general equivalence schemes, like in the case of the improved prenex normal form theorem we discuss in the next chapter. In the rest of this thesis, we will always write $\models^+_{\mathrm{G}}$, but reason in terms of sets of valuations according to the definition of $\models^+_{\mathrm{I}}$ whenever this helps making the argument more precise.

For closed formulas, i.e. *sentences*, we introduce the following abbreviation.

Definition 5.3.6 (satisfaction for sentences) *If φ is an IFG-*sentence *and $\mathfrak{A}$ a suitable model, we define:*

$$\mathfrak{A} \models^+_{\mathrm{G}} \varphi \quad (\text{“}\varphi \textit{ is true in } \mathfrak{A}\text{”}) \quad \stackrel{d}{\Leftrightarrow} \mathfrak{A} \models^+_{\mathrm{G}} \varphi[\{\lambda\}]$$
$$\mathfrak{A} \models^-_{\mathrm{G}} \varphi \quad (\text{“}\varphi \textit{ is false in } \mathfrak{A}\text{”}) \quad \stackrel{d}{\Leftrightarrow} \mathfrak{A} \models^-_{\mathrm{G}} \varphi[\{\lambda\}]$$

We will prove in the next chapter (corollary 6.2.3) that for sentences φ and for any set of valuations $V \subseteq A^X$: $\mathfrak{A} \models^{\pm}_{\mathrm{G}} \varphi[V]$ iff $\mathfrak{A} \models^{\pm}_{\mathrm{G}} \varphi$.

5.4 On the positive existential clause

In definition 5.3.3 we gave a positive existential clause that was formulated in terms of Y-saturation of subsets V_i of V. Because we require the V_i in the clause to form a *partition* (whereby they are disjoint), and we pick one domain element a_i for each V_i, it is not hard to see that this implicitly defines a Y-independent function $f : V \to A$. The clause $(\exists, +)$ of Definition 5.3.3 can thus easily be seen to have an alternative formulation making this function explicit:

Lemma 5.4.1 (alternative existential clause) $\mathfrak{A} \models^+_{\mathrm{I}} \exists x_{/Y}\varphi[V]$ *if and only if* $\mathfrak{A} \models^+_{\mathrm{I}} \varphi[V_{x:\,f}]$ *for some $f : V \to A$ that is independent of Y.*

The alternative formulation of the positive existential clause as in lemma 5.4.1, will be used in most of our arguments and proofs, as it combines well with our game theoretic intuitions.

As a side remark, we mention that if in the original clause of definition 5.3.3, we loosen the requirement of the V_i to be a *partition*, and instead let them just be a collection of subsets that together cover V, the clause can be seen to correspond to the *strategies as relations* approach, briefly mentioned at the end of section 2.4. This loosened version is equivalent to the partition-version by the axiom of choice, hence it would require the axiom of choice to prove the corresponding version of theorem 5.3.5. Choosing for the loosened version would avoid incorporating the axiom of choice in the inductive notion of satisfaction, and thereby make it possible to prove the conservative extension lemma in the next section without the axiom of choice. However, some of the results in the next chapter cannot (easily?) be generalized to a strategies-as-relations approach, so we stay with the strategy-as-function approach.

The inductive notion of satisfaction of definition 5.3.3 is based on Lemma 1.1 of [CK99]. However, we defined our existential clause differently, because we found a problem in the corresponding clause of that lemma. We will explain the problem, and how our alternative solves this problem.

Quote 5.4.2 (Lemma 1.1(e), [CK99, p. 21]) $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\varphi[V]$ *if and only if* $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V_x]$ *for some set* $V_x \subseteq V_{x\colon A}$ *that is independent of* Y *for* x.

The condition "V_x *is independent of* Y *for* x" is defined as follows ([CK99, p. 20]): whenever $v, w \in V_x$ coincide out of $Y \cup \{x\}$, then $v(x) = w(x)$. This is intended to correspond with the condition of Game Satisfaction that the strategy function $f_{\exists x_{/Y}\varphi}$ should be independent of Y.

This part of the lemma is incorrect at two points. First, we obviously need to add the condition that V_x is an x-variation of V, i.e. that V_x contains at least one x-variant of each $v \in V$ (cf. definition 5.3.4). Otherwise, any singleton $V_x \subseteq V_{x\colon A}$ satisfies the condition that V_x is Y-independent for x, thereby unintendedly making the inductive clause much weaker than game satisfaction. (Indeed, the condition that there is at least one x-variant for each $v \in V$, is needed to make f in the proof of this clause of [CK99] well-defined.)

But more subtly, the following example shows that in certain situations game satisfaction allows for signaling that is prohibited under the inductive clause of Lemma 1.1 (e).

Example 5.4.3: (Counterexample to [CK99], Lemma 1.1(e)) First, consider the formula $\psi := \exists x_{/y}[x = y]$ in a model $\mathfrak{A}$ with domain $A = \{0, 1\}$, and let V be the set of valuations $\{yz\colon 00, 11\}$. This situation could typically occur in the third stage of the game on the sentence $\forall y \exists z \exists x_{/y}[x = y]$.

It is easy to check that $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V]$: the function $f : V \to \{0, 1\}$ with $f(v) = v(z)$ is a winning strategy for Eloïse. The cited clause of Lemma 1.1(e) is also satisfied, by the set $V_x = \{xyz\colon 000, 111\}$. In both cases, *signaling* the value of y through the value of z is apparently allowed.

Now consider the same formula, in the same model, but with the slightly different set of valuations $V' = \{xy\colon 00, 11\}$. One could consider this situation as third stage of the game on the sentence $\forall y \exists x \exists x_{/y}[x = y]$.

Again, $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V']$ because $f' : V' \to \{0, 1\}$ with $f'(v) = v(x)$ is as valid a winning strategy for Eloïse as f in the previous case. On the other hand, the only x-variation of V' to satisfy the remaining atomic formula $(x = y)$ would be V' itself. But V' is *not* independent of y for x: the elements of V' coincide out of xy (where they are the empty valuation), which means that for independence they should assign the same value to x. They obviously don't.

Apparently the condition of Y-independence-for-x on the x-variation of V' is stricter than the condition of Y-independence on the function f'. $\diamond$

Note that an x-variation V_x of V that is independent of Y for x, contains *exactly one* x-variant for every $v \in V$. This means that in that case there is a function $f : V \to A$ such that

$$V_x = \{v(x\colon f(v))|v \in V\} = V_{x:\,f}$$

and from the Y-independence of V_x for x, we can deduce that this f is independent of Y. In the example above, we have seen that the converse is not necessarily true: there the set V' was also of the form

$$V'_{x:\,f'} = \{v(x\colon f'(v))|v \in V'\}$$

(with f' the winning strategy of Eloïse), but V' was not independent of Y for x.

In the example, we stumbled upon a subtle asymmetry between the definitions of independence for functions (as used in Game Satisfaction) and independence-for-x for sets of valuations (as used in Lemma 1.1(e) of [CK99]). The latter implied independence of the variable bound by the quantification, whereas the first doesn't do so. Note that this difference only occurs in situations where the quantified variable –x in the examples and definitions above– happens to be an element of the domain X of V.

To repair the difference, we had the choice to either include the implicit independence in a stricter notion of independence for functions, or to loosen up the notion of independence for sets of valuations. With the main argument that we like our language to be as general as possible with as little implicit conventions as possible, we chose the latter. This can easily be recognized in the formulation of clause $(\exists, +)$ in Lemma 5.4.1, which shows a direct correspondence of the condition in the inductive notion of satisfaction and the conditions in game satisfaction.

It is the equivalent, alternative clause of lemma 5.4.1 that we will mostly use, because of its direct correspondence to the application of choice functions. We also wanted to introduce the existential clause of definition 5.3.3, because it makes the definition of inductive satisfaction purely in terms of sets of valuations and Y-saturatedness: hence, the condition 'independence of Y' used for choice functions has its counterpart in 'Y-saturatedness' for sets of valuations. But from now on, we will use the clause that works best in our proofs and arguments.

In the rest of this thesis we will no longer distinguish between the two notions of satisfaction, as they were proved to coincide. The general notation we will use for satisfaction is $\models^{\pm}_{\mathrm{G}}$, which will be interpreted both in terms of strategies (mostly in the explanation of the examples) and inductively (in formal proofs).

5.5 Conservative extension of first order logic

As a first exercise for the semantics for open formulas defined in the previous section, we will prove explicitly that IFG-logic can be seen to be a conservative extension of first order logic. We have to explain what we mean by this, because IFG-formulas (whether or not their connectives and quantifiers are slashed), are evaluated with respect to a set of valuations, while first order formulas are evaluated only with respect to one valuation.

The reader should be aware that, as we explained in the introductory section of this chapter, it is not the case for IFG-formulas ψ in general that

$$\mathfrak{A} \models^{\pm}_{G} \psi[V] \text{ iff } \mathfrak{A} \models^{\pm}_{G} \psi[\{v\}] \text{ for all } v \in V \tag{5.1}$$

because the independence conditions in ψ will lose their meaning if we evaluate the formula with respect to only one valuation. E.g. $\mathfrak{A} \models^{+}_{G} \exists y_{/x}[x = y][\{x\colon 0\}]$ and $\mathfrak{A} \models^{+}_{G} \exists y_{/x}[x = y][\{x\colon 1\}]$, while *not* $\mathfrak{A} \models^{+}_{G} \exists y_{/x}[x = y][\{x\colon 0, 1\}]$. (Formulas that have the property (5.1) are called *flat* in [Hod97c].)

If an IFG-formula ψ contains no slashes, or more precisely: the sets of variables under the slashes in ψ are all empty, we say that ψ is a **first order formula**. We will see that (5.1) *does* hold for first order formulas. This follows from the next lemma, proving that our semantics for IFG-formulas is in a certain sense, a **conservative extension** of first order logic with classical semantics:

Lemma 5.5.1 ([CK99], Lemma 1.2, p. 22) *For every first order formula φ, structure $\mathfrak{A}$ and set of valuations $V \subseteq A^X$ with $Fv(\varphi) \subseteq X$:*

$$\mathfrak{A} \models^{+}_{G} \varphi[V] \textit{ iff } \mathfrak{A} \models \varphi[v] \textit{ for all } v \in V.$$

Similarly, $\mathfrak{A} \models^{-}_{G} \varphi[V]$ iff $\mathfrak{A} \not\models \varphi[v]$ for all $v \in V$.

Proof: First, we give the special cases of of definition 5.3.3 and lemma 5.4.1 in the case that $Y = \emptyset$ (which is the case for all quantifiers and connectives in first order formulas):

($\vee, +$) $\mathfrak{A} \models^{+}_{G} [\psi \vee \theta][V]$ iff $\mathfrak{A} \models^{+}_{G} \psi[V_1]$ and $\mathfrak{A} \models^{+}_{G} \theta[V_2]$, for some V_1 and V_2 such that $V = V_1 \cup V_2$

($\vee, -$) $\mathfrak{A} \models^{-}_{G} [\psi \vee \theta][V]$ iff $\mathfrak{A} \models^{-}_{G} \psi[V]$ and $\mathfrak{A} \models^{-}_{G} \theta[V]$.

($\exists, +$) $\mathfrak{A} \models^{+}_{G} \exists x\psi[V]$ iff there is an $f : V \to A$ such that $\mathfrak{A} \models^{+}_{G} \psi[V_{x\colon f}]$

($\exists, -$) $\mathfrak{A} \models^{-}_{G} \exists x\psi[V]$ iff $\mathfrak{A} \models^{-}_{G} \psi[V_{x\colon A}]$

We prove the lemma by induction on the structure of the first order formula φ:

(At) Immediate: by the definition of the atomic case in the definition of $\models^{\pm}_{G}$.

($\neg$) The following are equivalent:

- $\mathfrak{A} \models^{+}_{G} \neg\varphi[V]$
- $\mathfrak{A} \models^{-}_{G} \varphi[V]$ (by definition 5.3.3)
- $\mathfrak{A} \not\models \varphi[v]$ for all $v \in V$ (by induction hypothesis)
- $\mathfrak{A} \models \neg\varphi[v]$ for all $v \in V$.

Similarly (by the same arguments):

- $\mathfrak{A} \models_{\mathsf{G}}^{-} \neg\varphi[V]$
- $\mathfrak{A} \models_{\mathsf{G}}^{+} \varphi[V]$ (by definition 5.3.3)
- $\mathfrak{A} \models \varphi[v]$ for all $v \in V$ (by induction hypothesis)
- $\mathfrak{A} \not\models \neg\varphi[v]$ for all $v \in V$.

$(\vee, +)$ If $\mathfrak{A} \models_{\mathsf{G}}^{+} [\psi \vee \theta][V]$, let V_1, V_2 be such that $V = V_1 \cup V_2$ and $\mathfrak{A} \models_{\mathsf{G}}^{+} \psi[V_1]$ and $\mathfrak{A} \models_{\mathsf{G}}^{+} \theta[V_2]$. By induction hypothesis: $\mathfrak{A} \models \psi[v]$ for all $v \in V_1$ and $\mathfrak{A} \models \theta[v]$ for all $v \in V_2$. Hence: $\mathfrak{A} \models [\psi \vee \theta][v]$ for all $v \in V_1 \cup V_2 = V$.

Conversely, if $\mathfrak{A} \models [\psi \vee \theta][v]$ for all $v \in V$, we can let $V_1 := \{v \in V | \mathfrak{A} \models \psi[v]\}$, and $V_2 := \{v \in V | \mathfrak{A} \models \theta[v]\}$. Then, $V_1 \cup V_2 = V$, and by induction hypothesis: $\mathfrak{A} \models_{\mathsf{G}}^{+} \psi[V_1]$ and $\mathfrak{A} \models_{\mathsf{G}}^{+} \theta[V_2]$. Hence: $\mathfrak{A} \models_{\mathsf{G}}^{+} [\psi \vee \theta][V]$.

$(\vee, -)$ The following are equivalent:

- $\mathfrak{A} \models_{\mathsf{G}}^{-} [\psi \vee \theta][V]$
- $\mathfrak{A} \models_{\mathsf{G}}^{-} \psi[V]$ and $\mathfrak{A} \models_{\mathsf{G}}^{-} \theta[V]$ (by definition 5.3.3)
- $\mathfrak{A} \not\models \psi[v]$ and $\mathfrak{A} \not\models \theta[v]$, for all $v \in V$ (by induction hypothesis)
- $\mathfrak{A} \not\models [\psi \vee \theta][v]$ for all $v \in V$.

$(\exists, +)$ Suppose $\mathfrak{A} \models_{\mathsf{G}}^{+} \exists x\psi[V]$. Let $f : V \to A$ be a function such that $\mathfrak{A} \models_{\mathsf{G}}^{+} \psi[V_{x:\,f}]$. By induction hypothesis: $\mathfrak{A} \models \psi[v(x\colon f(v))]$ for all $v \in V$, but this implies $\mathfrak{A} \models \exists x\psi[v]$ for all $v \in V$.

Conversely, suppose $\mathfrak{A} \models \exists x\psi[v]$ for each $v \in V$. This means that for each v, there is an a such that $\mathfrak{A} \models \psi[v(x\colon a)]$. Due to the axiom of choice, there is then an $f : V \to A$ such that $\mathfrak{A} \models \psi[v(x\colon f(v))]$ for each $v \in V$. By induction hypothesis, it then follows that $\mathfrak{A} \models_{\mathsf{G}}^{+} \psi[V_{x:\,f}]$, hence $\mathfrak{A} \models_{\mathsf{G}}^{+} \exists x\psi[V]$.

$(\exists, -)$ The following are equivalent:

- $\mathfrak{A} \models_{\mathsf{G}}^{-} \exists x\psi[V]$
- $\mathfrak{A} \models_{\mathsf{G}}^{-} \psi[V_{x:\,A}]$ (by definition 5.3.3)
- $\mathfrak{A} \not\models \psi[v(x\colon a)]$ for all $v \in V$ and $a \in A$ (by induction hypothesis)
- $\mathfrak{A} \not\models \exists x\psi[v]$ for all $v \in V$.

$\triangleleft$

In particular, we have for **first order sentences** φ (cf. definition 5.3.6):

$$\mathfrak{A} \models_{\mathsf{G}}^{+} \varphi \quad \text{iff} \quad \mathfrak{A} \models \varphi$$
$$\mathfrak{A} \models_{\mathsf{G}}^{-} \varphi \quad \text{iff} \quad \mathfrak{A} \not\models \varphi$$

But more generally, it immediately follows from the previous lemma, that for an **arbitrary first order formula** ψ and a suitable valuation v for ψ in $\mathfrak{A}$:

$$\mathfrak{A} \models \psi[v] \text{ iff } \mathfrak{A} \models_{\mathsf{G}}^{+} \psi[\{v\}] \tag{5.2}$$
$$\mathfrak{A} \not\models \psi[v] \text{ iff } \mathfrak{A} \models_{\mathsf{G}}^{-} \psi[\{v\}] \tag{5.3}$$

(Note that these equivalences, in combination with the lemma, imply that (5.1) holds for first order formulas.) In this sense, IFG-logic and its semantics are a **conservative extension** of classical first order logic.

5.6 Signaling, IF-logic and regularity

In the previous section, using the inductive definition of satisfaction for IFG-formulas, we proved step-by-step that this notion of satisfaction coincides with Tarski semantics for first order formulas (in the sense of (5.2) and (5.3) above).

In this section we will return shortly to Hintikka's IF-language, which consisted of all first order sentences in negation normal form and the sentences resulting from them by application of the slash operator (according to some restricted rules, cf. definition 2.5.1). Because game theoretical semantics for first order sentences coincides with Tarski semantics, as noted at the end of section 2.4, it seems to follow easily that IF-logic is a conservative extension of first order logic (if we restrict the latter to sentences).

However, in chapter 3 we argued that Hintikka seems to take the syntactical procedure of Skolemization (in the usual outside-in order) as primitive rather than the game theoretical semantics he advocates. In section 3.9, we saw how this led Hintikka to introduce the *slashing convention*: the assumption that, in the semantic games, Eloïse always forgets her own previous moves. The introduction of this convention lets strategies in the semantics games consist of the intended Skolem functions (cf. also section 4.4), but it also turns semantic games for first order formulas into games of imperfect information (even: imperfect recall, cf. section 4.7). This means that it is no longer straightforward that the law of the excluded middle should hold, as the game-theoretical argument for this relies on the fact that these games are of *perfect* information. And if the law of the excluded middle would fail for first order sentences, this would compromise the conservative extension claim for IF-logic.

On the other hand, the combination of theorem 3.2.2 with Hintikka's approach 3.2.5 seems to show that, even under the slashing convention, a first order sentence is *classically* true in a suitable model ($\mathfrak{A} \models \varphi$) if and only if it is true under game theoretical semantics ($\mathfrak{A} \models\!\models_t \varphi$).

So, all seems to be well. But that we should never feel too safe too quickly, is demonstrated by the many examples given by Theo M.V. Janssen of unexpected *signaling* phenomena that surprise our classically trained minds (cf. e.g. [Jan02]). The value assigned to a 'forbidden' variable, may in some cases be signaled to a player by a strategically made earlier choice. The typical example of this phenomenon is Hodges' formula (3.10): $\forall x \exists z \exists y_{/x}[x = y]$. Eloïse may choose the value for z equal to x, and subsequently, the value for y equal to z. This way, she ensures $x = y$ without violating the condition that her strategy for y may not depend on x. (It is also problematic to interpret this epistemically –which may be accounted for by general conceptual problems with imperfect recall, cf. section 4.7: while Eloïse is supposed not to know the value of x, she apparently *does* know that the value of z is equal to the value of x.) Hintikka excluded the 'forbidden' dependencies in this example by the introduction of the slashing convention (cf. the quote on page 44 of this thesis).

But the slashing convention itself gives rise to new violations of our epistemic intuitions: how can Eloïse be supposed to forget all her own previous moves while she may be capable of remembering the earlier moves of her opponent? For example, compare the IF-formulas $\forall x \exists y[x = y]$ and $\exists x \exists y[x = y]$. In a semantic game for the first formula, Eloïse's winning strategy for the choice of y can (and must) use the value of x. However, she cannot copy this idea to get a winning strategy in a semantic game for the second formula. Making the implicit independence in the second formula explicit, it is equivalent to the IFG-formula: $\exists x \exists y_{/x}[x = y]$. In a game for that formula, she may not use the value of x. But she still *has* a winning strategy: she may coordinate her choice for x with her choice for y, by choosing them both by the same (constant) choice function.

We see that the slashing convention blocks the signaling possibility in formulas like Hodges' example, but that at the same time, it makes signaling an essential part of IF-logic. This is also demonstrated in the following similar, but a bit more complicated example:

$$\forall x \exists y[x = y \wedge \exists z[y = z]] \tag{5.4}$$

This first order sentence is classically true in all suitable models. As an IF-sentence, with the slashing convention, (5.4) translates into the IFG-sentence

$$\forall x \exists y[x = y \wedge \exists z_{/y}[y = z]] \tag{5.5}$$

In order to choose z equal to y, Eloïse may not use the value of y itself. However, she can use the value of x to signal the value of y, by the fact that she must have chosen y like that in order to satisfy the left conjunct $(x = y)$. By this 'trick', (5.4) turns out to be valid as an IF-sentence as well. But things go wrong if the crucial signal of x is blocked:

Example 5.6.1: (IF-logic may differ on first order sentences with nested quantification over the same variable) This example is a variation on formula (21) in [JD05]. Consider the first order sentence:

$$\forall x \exists y[x = y \wedge \forall x \exists z[y = z]]. \tag{5.6}$$

Like (5.4), this sentence is classically true in all models: the second quantification $\forall x$ is empty (i.e. the variable x does not occur in its scope), and may therefore be ignored.

If we read the first order sentence (5.6) as an IF-sentence, Hintikka's slashing convention applies to it (cf. section 3.9). This means that the inner existential quantification $\exists z$ is taken to be implicitly independent of the outer existential quantification $\exists y$. We make this implicit independence explicit in the following IF*G*-formula, that thereby has the same meaning as the IF-sentence (5.6):

$$\forall x \exists y[x = y \wedge \forall x \exists z_{/y}[y = z]]. \tag{5.7}$$

In comparison with (5.5), Eloïse now has a problem: even if the variable x does not occur in $\exists z_{/y}[y = z]$, Abélard did choose a possibly different value for x before Eloïse gets to choose z. So, she can no longer trust the value of x as a signal for the value of y, hence: she has no winning strategy!

This is supported by a step-by-step analysis of the IFG-sentence by the inductive definition of satisfaction. Consider this sentence in a model $\mathfrak{A}$ with domain $\{0,1\}$. We then have:

$$\begin{array}{lcr}
\mathfrak{A} \models^+_{\mathsf{G}} & \forall x \exists y[x = y \wedge \forall x \exists z_{/y}[y = z]], & \text{iff} \\
\mathfrak{A} \models^+_{\mathsf{G}} & \forall x \exists y[x = y \wedge \forall x \exists z_{/y}[y = z]][\{\lambda\}], & \text{iff} \\
\mathfrak{A} \models^+_{\mathsf{G}} & \exists y[x = y \wedge \forall x \exists z_{/y}[y = z]][\{x\colon 0,1\}], & \text{iff} \\
\mathfrak{A} \models^+_{\mathsf{G}} & (x = y) \wedge \forall x \exists z_{/y}[y = z][\{x\colon 0,1\}_{y\colon f}] &
\end{array}$$

for some $f : \{x\colon 0,1\} \to \{0,1\}$. In order to let the y-variants of $(x\colon 0)$ and $(x\colon 1)$ both to satisfy the left conjunct $(x = y)$, there is no other option than to let $f(x\colon 0) = 0$ and $f(x\colon 1) = 1$. We then need to verify:

$$\begin{array}{lcr}
\mathfrak{A} \models^+_{\mathsf{G}} & \forall x \exists z_{/y}[y = z][\{xy\colon 00,11\}], & \text{hence if} \\
\mathfrak{A} \models^+_{\mathsf{G}} & \exists z_{/y}[y = z][\{xy\colon 00,10,01,11\}] &
\end{array}$$

In order to verify the latter, we need a function $g : \{xy\colon 00,10,01,11\} \to \{0,1\}$ that is constant on the equivalence classes $\{xy\colon 00,01\}$ and $\{xy\colon 10,11\}$ of the relation $\sim_y$. But let's say that $g(v) = c$ for both elements in the first set, then one of the z-variants in $\{xyz\colon 00c, 01c\}$ fails to satisfy the atomic formula $(y = z)$. (Similarly for the second equivalence class.) Hence:

$$\mathfrak{A} \not\models^+_{\mathsf{G}} \forall x \exists y[x = y \wedge \forall x \exists z_{/y}[y = z]].$$

To see the problem from a different angle, let's look at the truth condition of the first order sentence (5.6) by Skolemization:

$$\exists f \exists g \forall x[x = f(x) \wedge \forall x[f(x) = g(x)]].$$

We observe that the argument x of the Skolem function f falls under two different quantifications over x. This makes the replacement of y within the scope of the inner quantification over x *unsound*: the occurrences of y would always get the same value assigned, while the term $f(x)$, by which y is replaced, may have different values at different places in the formula. We conclude from this example that apparently, an implicit precondition for Skolemization is that no nested quantification over the same variable occurs in the formula. ◇

We conclude that the evaluation of unslashed IF-sentences (i.e. first order sentences with the slashing convention) is not always the same as their classical evaluation as first order formula. In fact, we have shown an unslashed IF-sentence that is *undecided*. Apparently, the slashing convention can let the law of the excluded middle fail for first order sentences. It all depends on the occurrence of nested quantification over the same variable. We will call this syntactic phenomenon *irregularity* (and its opposite *regularity*). We will encounter it many times in the next chapter as the cause of more problems.

With regular first order sentences, problems as in the example above do not occur: we can then safely apply the Skolemization procedure and use the argument that we gave at the end of the second paragraph of this section. So:

Theorem 5.6.2 (conservative extension theorem for IF-logic) *IF-logic is a conservative extension of the classical evaluation for* regular *first order sentences.*

Proof: For the Skolemization procedure to be sound, we need the first order sentences to be regular. The theorem then follows from theorem 3.2.2 and Hintikka's approach 3.2.5. ◁

In [HA59, p. 74], Hilbert and Ackermann defined the language of first order predicate logic in such a way that the formulas never contain nested quantification over the same variable (nor do variables occur both free and bound in one formula).[1] Nowadays, it is rather common to define the first order language without these restrictions. Our example above indicates that for the interpretation of first order sentences as IF-sentences, it is important to adhere to these restrictions.

5.7 Conclusions

In this chapter, we have introduced and explained the game semantics for open IFG-formulas as given by [CK99]. We also gave an alternative, inductive, definition of satisfaction for IFG-formulas that is entirely in terms of sets of valuations. We proved the two notion of satisfaction to be equivalent. Moreover, we gave an alternative existential clause for the inductive definition of satisfaction, in terms of functions, equivalent to the one we gave in terms of sets of valuations. Because of its close conceptual correspondence with the game theoretic terminology we use informally, it is this clause that we will mostly use.

We used the semantics for open IFG-formulas to prove how IFG-logic is a conservative extension of first order logic.

Our inductive definition of satisfaction was based on [CK99, Lemma 1.1]. However, we demonstrated a problem with the original existential clause. This problem occurred if the variable bound by the existential quantifier, was already in the domain X of the set of valuations with which the formula was evaluated. We also demonstrated how the slashing convention challenges the claim that IF-logic is a conservative extension of classical logic: if an unslashed IF-sentence contains nested quantification over the same variable, signals that are crucial to overcome Eloïse's imperfect recall (induced by the slashing convention), may be lost. IF-logic therefore only extends the *regular* first order sentences: sentences in which nested quantification over the same variable does not occur.

Similar situations cause problems for the steps leading to the prenex normal form theorem of [CK99]. We will show this in the next chapter, where we also introduce the solutions to recover the prenex normal form theorem for regular IFG-sentences.

[1] Our attention was drawn to this fact at a course taught by Saul Kripke in Utrecht in January 2001, which started with a long and energetic defence of what Kripke referred to as the *Hilbert-Ackermann conventions.*

Chapter 6

The prenex normal form theorem

In this chapter, we critically study the equivalence schemes of [CK99], used to prove a prenex normal form theorem for IFG-formulas. The proof of this result goes along the same type of equivalence schemes as in classical first order logic: renaming of bound variables, quantifier extraction, and substitution of equivalents. Lifting these schemes to logic with imperfect information turns out to be even less straightforward than described in [CK99]. Independence conditions make the formulas sensitive for signaling phenomena. In particular, nested quantification over the same variable and related situations are shown to cause problems, by the fact that in those cases signaling can be blocked unexpectedly.

We show that the notion of equivalence used in the equivalence schemes, is too strict for renaming of bound variables. We prove a couple of general lemmas, refine definitions and sharpen preconditions, allowing us to restore (a restricted version of) the prenex normal form theorem. We also prove how slashed connectives can be eliminated. All these results are proved for a *strong* notion of equivalence, and the arguments stay within the language and semantics of IFG-logic, i.e. without using translations to other, 'meta'-systems (like Σ_1^1).

The results presented in this chapter are joint work with Xavier Caicedo and Theo M.V. Janssen ([CDJ]).

6.1 Introduction

Recall that a formula is in prenex form if it consists of a (possibly empty) sequence of quantifiers, followed by a *quantifier-free* formula ψ:

$$Q_2x_1Q_2x_2\dots Q_kx_k\psi(x_1,x_2,\dots,x_k).$$

It is a well known fact that every first order formula can be transformed into an equivalent formula in prenex form. We call this result the *prenex normal form theorem* for first order logic.

The prenex normal form theorem for IFG-formulas presented in [CK99], generalizes the first order prenex normal form theorem. It is obtained following the same steps: if necessary, quantified variables are renamed, then quantifiers are 'pulled over connectives' to move them to the front of the formula. All steps are proved for IFG-formulas in terms of the semantics presented in the previous chapter. (This means that our prenex normal form theorem does not depend on a similar result for some classical logic, as would for example be the case if we would transform the sentences into Σ_1^1-sentences, transform these into prenex form, and then translate the result back into slashed sentences.) Moreover, the prenex normal form theorem for IFG-formulas is valid in a strong sense of equivalence (which is, for example, not the case for the generalized Skolem form mentioned in [Hin96, p. 60]).

But as the list of counterexamples in this chapter indicates, the original result of [CK99] was not immune for the traps of *signaling*, that always lurk in the background of our logic with imperfect information. In section 5.6, we described signaling as a situation in which information is conveyed, that was not (supposed to be) available to a player in a game. With signaling in IFG-logic, we mean situations in which the imperfect information requirements (induced by the slash operator) can be overcome through values assigned to other variables. These may tell a player something about the value of the variable she is not allowed to use (or at least: not explicitly).

We discover situations in which signaling either becomes possible or blocked by subtle changes in either the IFG-formula or the domain of the valuations with which it is evaluated. These situations make it necessary to reformulate a number of the results from [CK99], and inspire us to formulate and prove some new insights. Most of the problematic situations were presented (but not yet solved) in the paper [JD05]. Solving them turned out to be far from straightforward, as demonstrated in the course of this chapter.

Histories vs valuations

To clarify the nature of the problems, we point out that they are a consequence of the fact that we model the semantic games using *valuations*. In the logical context that we are working in, this is the natural option. But when we stated in chapter 4 that there is a correspondence between *histories* and pairs of subformulas with *valuations*, we implicitly used the assumption made on page 57 that no nested quantification over the same variable occurs in the sentence for which the game is modeled.

We needed the assumption that the sentence is regular, because there is a difference between histories and valuations that occurs if we play semantic games for *irregular* formulas: where in a history every choice for a domain element is stored at the 'chronological' point where the choice was made (with no reference to the quantified variable giving rise to the move), in a valuation it is stored as the assignment for the specific bound variable. This means that, if in one run of a game a second value must be chosen for the same variable, the first value chosen for that variable will be overwritten in the *valuation*, while in the corresponding *history*, the two values are still available (we could refer to the first and the second value chosen for the variable).

Working with open formulas, the same type of problem can also occur if a formula contains one variable both free and bound, or even if it is evaluated with respect to valuations whose domain contain variables that occur bound in the formula. We will show how these technical problems can be solved.

A short comparison with trump semantics

We remark that most counterexamples and results in this chapter can easily be transferred (in some form) to Hodges' trump semantics (as defined in [Hod97a]). The main difference between trump semantics and the semantics of [CK99] (which we presented in the previous chapter), is that Hodges acknowledges the non-trivial role of the domains of the valuations by considering formulas only *in combination with* a fixed set of variables. These variables are all to be treated as free variables of the formula, even if they do not occur in the formula itself. The *trumps* of a formula(-combination) $\psi(X)$ in a model $\mathfrak{A}$ are the sets $V \subseteq A^X$ for which $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V]$ (in our terminology). The set of trumps of a formula(-combination) $\psi(X)$ in a model constitute its *meaning* in that model. An immediate consequence is that a formula $\psi(x)$ has a different meaning than a formula $\psi(x, z)$ even if z does not occur in ψ; that this is not so strange in the context of logic with imperfect information may be demonstrated by the comparison of the sentences $\forall x[\psi(x)]$ and $\forall x \exists z[\psi(x, z)]$ where $\psi = \exists y_{/x}[x = y]$ (the latter sentence is then Hodges' formula (3.10)).

Unfortunately, we have not yet finished a full account of how the results in this chapter translate to trump semantics, but in some remarks in the next section we indicate the noted relations to Hodges' work. The next section proves results that in Hodges' terms are of the form: "If V is a trump for the formula $\varphi(X)$, then V' is a trump for the formula $\varphi'(X')$".

6.2 Some monotonicity results

For later use, we will now state some facts on satisfaction for IFG-formulas with respect to different sets of valuations: we prove that taking subsets preserves satisfaction, and that under certain circumstances the domain of the valuations can be extended by new variables. This corresponds to proposition 2 of [Hod97c].

Lemma 6.2.1 (Downward monotonicity) *Let φ be an IFG-formula, and $\mathfrak{A}$ a suitable model for φ. Then for any set of valuations $V \subseteq A^X$ with $Fv(\varphi) \subseteq X$, and for any $W \subseteq V$: $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V]$ implies $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[W]$.*

Proof: We use induction on the complexity of φ. The atomic case and induction for $\neg$ are immediate.

$(\vee, +)$ $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi_1 \vee_{/Y} \varphi_2)[V]$ means the existence of subsets V_1 and V_2 of V, both Y-saturated in V with $V_1 \cup V_2 = V$, such that $\mathfrak{A} \models^+_{\mathsf{G}} \varphi_i[V_i]$. Define $W_i :=$

$V_i \cap W$ ($i = 1, 2$). Then the W_i are Y-saturated in W, $W_1 \cup W_2 = W$, and by induction hypothesis: $\mathfrak{A} \models^+_{\mathsf{G}} \varphi_i[W_i]$ for $i = 1, 2$. Hence: $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi_1 \vee_{/Y} \varphi_2)[W]$

$(\vee, -)$ $\mathfrak{A} \models^-_{\mathsf{G}} (\varphi_1 \vee_{/Y} \varphi_2)[V]$ means $\mathfrak{A} \models^-_{\mathsf{G}} \varphi_i[V]$ for $i = 1, 2$. By induction hypothesis this implies $\mathfrak{A} \models^-_{\mathsf{G}} \varphi_i[W]$, hence: $\mathfrak{A} \models^-_{\mathsf{G}} (\varphi_1 \vee_{/Y} \varphi_2)[W]$.

$(\exists, +)$ $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\varphi[V]$ means $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V_{x:\, f}]$ for some Y-independent $f : V \to A$. Obviously, the restriction of f to W is also independent of Y, and by induction: $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[W_{x:\, f}]$. Thus: $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\varphi[W]$.

$(\exists, -)$ $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\varphi[V]$ means $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[V_{x:\, A}]$. By induction, this implies $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[W_{x:\, A}]$, and hence: $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\varphi[W]$.

$\lhd$

Note that, of course, "upward monotonicity" is more problematic. For example, if $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[W]$ for some $V, W \subseteq A^X$ with $Fv(\varphi) \subseteq X$, then not necessarily $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V \cup W]$. (Take for example: $\varphi = \exists y_{/x}[x = y]$, $V = \{x\colon 0\}$, and $W = \{x\colon 1\}$.)

From the previous lemma, it follows that the sets V_1, V_2 in the clause for disjunction of inductive satisfaction (item $(\vee, +)$ of definition 5.3.3) may be taken to be disjoint. (To check Y-saturatedness: the intersection of two Y-saturated sets V_1, V_2 is easily seen to be Y-saturated in V, and therefore: $V_i \setminus (V_1 \cap V_2)$ is also Y-saturated in V. So, if we have Y-saturated V_i in V with non-empty intersection, we can equivalently take the disjoint sets $V_1 \setminus (V_1 \cap V_2)$ and V_2.)

An implicit parameter in the evaluation of a formula with a set of valuations $V \subseteq A^X$, is the domain X of the valuations. (This is a difference with trump semantics, where the domains are explicit parameters with the formula.) In IFG-logic, the values assigned to variables that do not occur in the formula could make a difference in the evaluation. For example, take $\mathfrak{A}$ a model with domain $A = \{0, 1\}$, let φ be the formula $\exists y_{/z}[z = y]$, and let $V := \{z\colon 0, 1\}$ and $V_x := \{xz\colon 00, 11\}$. The valuations in V_x assign the same values to the variables occurring in φ (only z in this case), but $\mathfrak{A} \not\models^+_{\mathsf{G}} \varphi[V]$, while $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V_x]$. The values assigned to x in V_x make that the two valuations in V_x do not coincide out of $\{z\}$, hence make the independence condition vacuous for this situation.

We now prove two general lemmas that compare the evaluation of a formula with respect to a given set of valuations V, and sets of extensions of these valuations to a larger domain of variables. First we show that the evaluation will not change as long as we extend all the valuations in V with the same combinations of values for the new variables, i.e. if we take a Cartesian product (this result seems to be a more general case of Proposition 3 of [Hod97c, p. 57]):

Lemma 6.2.2 (Expansion by Cartesian products) *If φ is an IFG-formula, and V, W sets of valuations that satisfy*

1. $V \subseteq A^X$ with $Fv(\varphi) \subseteq X$, and

2. $\emptyset \neq W \subseteq A^Z$ *with* $Z \cap X = \emptyset$

then:

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V] \textit{ if and only if } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V \times W]$$

Proof: We prove the lemma by induction on the complexity of φ. To start the induction: let φ be atomic, and V, W sets of valuations satisfying conditions 1-2 for φ. It classically holds for all $v \in V, w \in W$ that $\mathfrak{A} \models \varphi[v]$ if and only if $\mathfrak{A} \models \varphi[(v, w)]$, and $\mathfrak{A} \not\models \varphi[v]$ if and only if $\mathfrak{A} \not\models \varphi[(v, w)]$. It follows by definition 5.3.3 and theorem 5.3.5 that: $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V]$ iff $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V \times W]$. Note that from right to left we need that $W \neq \emptyset$ (otherwise $V \times W$ would be empty, hence $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V \times W]$ regardless of V).

For the induction step, assume that φ is non-atomic, that V, W are sets of valuations satisfying the condition 1-2 for φ, and that for all ψ with lower complexity than φ, we have proved:

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \psi[V'] \text{ if and only if } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \psi[V' \times W']$$

for any $V' \subseteq A^{X'}, W' \subseteq A^{Z'}$ that satisfy the conditions 1-2 with respect to ψ (i.e. $Fv(\psi) \subseteq X'$, $W' \neq \emptyset$ and $X' \cap Z' = \emptyset$). We prove that

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V] \text{ iff } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V \times W]$$

by distinction of the following cases:

($\neg$) If $\varphi = \neg\psi$, the result for φ follows immediately from the induction hypothesis applied to ψ with V, W.

($\vee$) Suppose $\varphi = \psi_1 \vee_{/Y} \psi_2$.

(+) From left to right: suppose that $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V]$. We can then determine $V_1, V_2 \subseteq V$ with $V_1 \cup V_2 = V$, both Y-saturated in V, such that $\mathfrak{A} \models^{+}_{\mathsf{G}} \psi_i[V_i]$ for $i = 1, 2$.. Then the sets $V_1 \times W, V_2 \times W$ are Y-saturated in $V \times W$, $(V_1 \times W) \cup (V_2 \times W) = V \times W$, and by induction hypothesis, $\mathfrak{A} \models^{+}_{\mathsf{G}} \psi_i[V_i \times W]$ for both i. Hence: $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V \times W]$.

To prove the converse, suppose $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V \times W]$. We pick some $w \in W$ ($W \neq \emptyset$), then by lemma 6.2.1: $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V \times \{w\}]$ We can then determine $V_1, V_2 \subseteq V$ such that $\mathfrak{A} \models^{+}_{\mathsf{G}} \psi_1[V_1 \times \{w\}]$ and $\mathfrak{A} \models^{+}_{\mathsf{G}} \psi_2[V_2 \times \{w\}]$ with $V_i \times \{w\}$ Y-saturated in $V \times \{w\}$ for both i and $V_1 \cup V_2 = V$. V_1, V_2 are clearly Y-saturated in V, and $\mathfrak{A} \models^{+}_{\mathsf{G}} \psi_i[V_i]$ for both i by induction hypothesis. Hence: $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V]$.

(-) This follows immediately from the induction hypothesis applied to ψ_1 and ψ_2 with V, W.

($\exists$) Suppose $\varphi = \exists x_{/Y}\psi$. In the evaluation of φ, a value will be assigned to the variable x. We have not excluded (and do not want to exclude) the possibility that x is in the domain Z of W. In order to make sure that we can validly apply the induction hypothesis, we will use the set $W_{-x} \subseteq A^{Z-\{x\}}$ of the restrictions of the elements of W to all variables in Z except x: $W_{-x} := \{w \restriction (Z - \{x\}) | w \in W\}$. Note that if $x \notin Z$, we have $W_{-x} = W$.

(+) From left to right: $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\psi[V]$ implies $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{x:\,f}]$ for some Y-independent $f : V \to A$, and by induction hypothesis: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{x:\,f} \times W_{-x}]$; we use W_{-x} instead of W, because x might be in $Z = dom(W)$, in which case the domains of $V_{x:\,f}$ and W violate the requirement of being disjoint. Note that $V_{x:\,f} \times W_{-x} = (V \times W)_{x:\,g}$ for the function $g : V \times W \to A$ defined by $g(v, w) = f(v)$. So: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[(V \times W)_{x:\,g}]$, and because g inherits Y-independence from f: $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\psi[V \times W]$.
Conversely, if $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\psi[V \times W]$, then by lemma 6.2.1, for any arbitrary $w \in W$: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V \times \{w\}]$. Pick a $w \in W$ ($W \neq \emptyset$), and let $f : V \times \{w\} \to A$ be a Y-independent function such that $\mathfrak{A} \models^+_{\mathsf{G}} \psi[(V \times \{w\})_{x:\,f}]$. Define $g : V \to A$ by $g(v) = f(v, w)$ and let $w_{-x} := w \restriction (Z - \{x\})$. Then: $(V \times \{w\})_{x:\,f} = V_{x:\,g} \times \{w_{-x}\}$, hence: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{x:\,g} \times \{w_{-x}\}]$. Applying the induction hypothesis, we may infer that $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{x:\,g}]$, and because g inherits Y-independence from f: $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\psi[V]$.

(-) The following statements are equivalent:

* $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\psi[V]$
* $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V_{x:\,A}]$
* $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V_{x:\,A} \times W_{-x}]$ (by induction hypothesis)
* $\mathfrak{A} \models^-_{\mathsf{G}} \psi[(V \times W)_{x:\,A}]$ (because $(V \times W)_{x:\,A} = V_{x:\,A} \times W_{-x}$)
* $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\psi[V \times W]$.

◁

This lemma has a reassuring and important corollary for sentences φ, justifying the abbreviation we introduced in definition 5.3.6:

Corollary 6.2.3 (Evaluation of sentences) *Given an IFG*-sentence *φ, and a suitable model $\mathfrak{A}$, then for any non-empty sets of valuations $V \subseteq A^X$ of some set of variables X in $\mathfrak{A}$:*

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V] \textit{ if and only if } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[\{\lambda\}]$$

Proof: Let V and X play the role of W and Z in the previous lemma. ◁

Lemma 6.2.2 implies that if x does not occur in φ nor in $X = dom(V)$, and $V_x \subseteq V_{x:\,A}$ is an x-variation of V of the form $V \times W$ with $W \subseteq A^{\{x\}}$, then

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V] \text{ if and only if } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V_x].$$

This observation will be useful in several of the proofs later in the chapter. The situation of our example on page 108, just before lemma 6.2.2, was different however:

it showed us that values assigned to variables that do not occur in the formula could influence its evaluation (they could *signal* the values of forbidden variables to the player).

The next result shows that if V_x is an *arbitrary* x-variation of $V \subseteq A^X$ (with x not occurring in φ or X), then we should make the quantifiers and connectives in φ independent of x to guarantee the same evaluation. It corresponds to Proposition 8 in [Hod97c]. We will use the following notation:

Definition 6.2.4 (Slashed formulas) *For any IFG-formula φ and any variable x that does not occur bound in φ, the formula $\varphi_{/x}$ is defined inductively as follows:*

(At) for atomic φ: $\varphi_{/x} := \varphi$

($\neg$) if $\varphi = \neg\varphi'$, then $\varphi_{/x} := \neg(\varphi'_{/x})$

($\vee$) if $\varphi = \varphi_1 \vee_{/Y} \varphi_2$, then $\varphi_{/x} := (\varphi_{1/x}) \vee_{/Y,x} (\varphi_{2/x})$

($\exists$) if $\varphi = \exists z_{/Y}\varphi'$, then $\varphi_{/x} := \exists z_{/Y,x}(\varphi'_{/x})$

Lemma 6.2.5 (Adding mute variables) *For all IFG-formulas φ: if $V \subseteq A^X$ is a set of valuations for φ, and x is a variable that does not occur in φ or X, then for any x-variation V_x of V (cf. definition 5.3.4):*

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V] \textit{ if and only if } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi_{/x}[V_x]$$

Proof: We use induction on the complexity of φ. The atomic case is immediate. So, let φ be a non-atomic IFG-formula, $V \subseteq A^X$ a set of valuations for φ, and x a variable that does not occur in φ or X. Also, suppose we have proved the lemma for all IFG-formulas of lower complexity than φ.

For $v \in V$, let v_x denote an arbitrary x-variant of v. Observe that because $x \notin X$, we have for all $v, w \in V$: $v_x \sim_{Y,x} w_x$ if and only if $v \sim_Y w$.

($\neg$) Suppose $\varphi \equiv \neg\psi$. This case follows easily from the induction hypothesis.

($\vee$) Suppose $\varphi \equiv \psi_1 \vee_{/Y} \psi_2$.

(+) From the observation above, it follows easily that if $V_i \subseteq V$ is Y-saturated in V, then $(V_i)_{x:\,A} := \{v_x \in V_x | v \in V_i\}$ is $(Y \cup \{x\})$-saturated in V_x, and if $V_1 \cup V_2 = V$, then also $(V_1)_{x:\,A} \cup (V_2)_{x:\,A} = V_x$. These observations together with the induction hypothesis prove that $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V]$ implies $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)_{/x}[V_x]$
Conversely, if $W_i \subseteq V_x$ is $(Y \cup \{x\})$-saturated in V_x, then $(W_i)_X := \{v \in V | v_x \in W_i\}$ is Y-saturated in V. Also, if $W_1 \cup W_2 = V_x$, then $(W_1)_X \cup (W_2)_X = V$. These observations together with the induction hypothesis prove that $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)_{/x}[V_x]$ implies $\mathfrak{A} \models^{+}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V]$

(-) $\mathfrak{A} \models^{-}_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V]$ if and only if $\mathfrak{A} \models^{-}_{\mathsf{G}} \psi_i[V]$ $(i = 1, 2)$, if and only if $\mathfrak{A} \models^{-}_{\mathsf{G}} \psi_{i/x}[V_x]$ (by induction hypothesis) if and only if $\mathfrak{A} \models^{-}_{\mathsf{G}} (\psi_{1/x} \vee_{/Y,x} \psi_{2/x})[V_x]$

($\exists$) Suppose $\varphi \equiv \exists u_{/Y}\psi$.

(+) Suppose $\mathfrak{A} \models^+_{\mathsf{G}} \exists u_{/Y}\psi[V]$, then $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{u:\,f}]$ for some Y-independent $f : V \to A$. Let $g : V_x \to A$ be defined by $g(v_x) = f(v)$, then g is independent of $Y \cup \{x\}$. Note that the set $(V_x)_{u:\,g}$ is an x-variation of $V_{u:\,f}$, so that we can apply the induction hypothesis to get $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_x)_{u:\,g}]$. Then, by $(Y \cup \{x\})$-independence of g: $\mathfrak{A} \models^+_{\mathsf{G}} (\exists u_{/Y}\psi)_{/x}[V_x]$
Conversely, if $\mathfrak{A} \models^+_{\mathsf{G}} (\exists u_{/Y}\psi)_{/x}[V_x]$, then $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_x)_{u:\,g}]$ for some $g : V_x \to A$ independent of $Y \cup \{x\}$. Note that if $v \in V$, then g assigns the same value to all $v_x \in V_x$. This makes $f : V \to A$ with $f(v) = g(v_x)$ for any $v_x \in V_x$ a well-defined function, that is independent of Y. Note that the set $(V_x)_{u:\,g}$ is again an x-variation of $V_{u:\,f}$ so that we can apply the induction hypothesis to get $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{u:\,f}]$, from which we conclude: $\mathfrak{A} \models^+_{\mathsf{G}} \exists u_{/Y}\psi[V]$.

(-) $\mathfrak{A} \models^-_{\mathsf{G}} \exists u_{/Y}\psi[V]$ if and only if $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V^u]$, if and only if $\mathfrak{A} \models^-_{\mathsf{G}} \psi_{/x}[(V^u)_x]$ (by induction hypothesis). Any x-variation $(V^u)_x$ of V^u can be seen as the set $(V_x)^u$ for some x-variation V_x of V, and vice versa. So, the latter is equivalent to $\mathfrak{A} \models^-_{\mathsf{G}} \exists u_{/Y,x}\psi_{/x}[V_x]$.

$\triangleleft$

In Example 2.4 of [CK99, p. 25], it is shown that even if the variable x does not occur in φ, the formula $\exists x\varphi$ can have a different evaluation than φ: consider $\varphi = \exists z_{/y}[z = y]$, a model $\mathfrak{A}$ with domain $\{0, 1\}$, and $V = \{z\colon 0, 1\}$; then $\mathfrak{A} \models^+_{\mathsf{G}} \exists x\varphi$ while $\mathfrak{A} \not\models^+_{\mathsf{G}} \varphi$. (Note that this is a reformulation in terms of open formulas of Hodges' example, as it corresponds to the fact that $\forall y\exists z_{/y}[y = z]$ is not equivalent to $\forall y\exists x\exists z_{/y}[y = z]$.) A consequence of the previous lemma is that we *do* have that φ always has the same evaluation as $\exists x\varphi_{/x}$.

We will prove a counterpart to this result, in the sense that it puts independence conditions on the added quantifier instead of on the formula. This counterpart is a corollary (corollary 6.2.7) to the following lemma:

Lemma 6.2.6 (Monotonicity under x-variations) *For any IFG-formula φ, and suitable model $\mathfrak{A}$: if $V \subseteq A^X$ is a set of suitable valuations, x a variable that does not occur in φ nor in X, and Z the set of all free variables of φ that occur under the slashes in φ; then for any Z-independent function $f : V \to A$:*

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V] \textit{ iff } \mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V_{x:\,f}]$$

Proof: By induction on the structure of φ. The start of the induction, for atomic φ, is immediate. So, suppose φ is non-atomic, and that for all ψ of lower complexity, we have proved that if $W \subseteq A^Y$ is a set of valuations for ψ in $\mathfrak{A}$, y a variable that does not occur in ψ nor in Y, and $g : W \to A$ a Z'-independent function (with Z' the set of free variables of ψ occurring under the slashes), then $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \psi[W]$ if and only if $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \psi[W_{y:\,g}]$.

Let $\mathfrak{A}$ be a suitable model, $V \subseteq A^X$, x a variable that does not occur in φ nor in X, and $f : V \to A$ a function that is independent of Z, the set of free variables of φ that occur under the slashes. We distinguish the following cases:

$(\neg, \pm)$ If $\varphi \equiv \neg\psi$, then the claim for φ follows immediately from the induction hypothesis.

$(\vee)$ If $\varphi \equiv (\psi_1 \vee_{/Y} \psi_2)$:

$(+)$ From left to right: suppose $\mathfrak{A} \models^+_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V]$, so pick $V_1, V_2 \subseteq V$ that are Y-saturated in V with $\mathfrak{A} \models^+_{\mathsf{G}} \psi_i[V_i]$. By the induction hypothesis, the latter is equivalent to $\mathfrak{A} \models^+_{\mathsf{G}} \psi_i[(V_i)_{x:\,f}]$ (which is applicable because x does not occur in X, ψ_1 or ψ_2; for f, independence of Z implies independence of the free variables under the slashes in the subformulas ψ_1 and ψ_2). As we will prove later, the sets $(V_i)_{x:\,f}$ are Y-saturated in $V_{x:\,f}$ and $V_{x:\,f} = (V_1)_{x:\,f} \cup (V_2)_{x:\,f}$, so: $\mathfrak{A} \models^+_{\mathsf{G}} (\psi_1 \vee_{/Y} \psi_2)[V_{x:\,f}]$.
The argument from right to left is similar: we use that by definition of $V_{x:\,f}$, any subset $(V_{x:\,f})_i$ of $V_{x:\,f}$ is of the form $(V_i)_{x:\,f}$ for some $V_i \subseteq V$. This V_i is unique because $x \notin X$.
We will now prove that for every $W \subseteq V$: W is Y-saturated in V if and only if $W_{x:\,f}$ is Y-saturated in $V_{x:\,f}$, in other words: $W \subseteq V$ is closed under the relation $\sim_Y$ on V if and only if $W_{x:\,f} \subseteq V_{x:\,f}$ is closed under the relation $\sim_Y$ on $V_{x:\,f}$. This follows easily from the fact that $v \sim_Y w$ iff $v(x\colon f(v)) \sim_Y w(x\colon f(w))$: if $v, w \in V$ coincide out of Y, then they also coincide out of Z (because $Y \subseteq Z$ by definition of Z), so by Z-independence of f: $f(v) = f(w)$, and hence: $v(x\colon f(v))$ and $w(x\colon f(w))$ coincide out of Y; the converse follows from $x \notin X$.

(-) This follows directly from the induction hypothesis applied to the ψ_i and (with V, x, and f as for φ).

$(\exists)$ Suppose $\varphi \equiv \exists z_{/Y}\psi$. We will use the function $f^* : V_{z:\,A} \to A$ defined by $f^*(v(z\colon a)) := f(v)$ for all $v \in V, a \in A$ (then f^* is independent of $Z \cup \{z\}$).

$(+)$ If $\mathfrak{A} \models^+_{\mathsf{G}} \exists z_{/Y}\psi[V]$, pick a Y-independent function $g : V \to A$ such that $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{z:\,g}]$. By induction $\mathfrak{A} \models^+_{\mathsf{G}} \psi[(V_{z:\,g})_{x:\,f^*}]$. It is easily verified (because x and z are distinct by the assumption that x does not occur in φ) that $(V_{z:\,g})_{x:\,f^*} = (V_{x:\,f})_{z:\,g^*}$ with $g^* : V_{x:\,A} \to A$ defined by $g^*(v(x\colon a)) := g(v)$ for all $v \in V, a \in A$. Therefore: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[(V_{x:\,f})_{z:\,g^*}]$, and hence (because g^* inherits Y-independence from g): $\mathfrak{A} \models^+_{\mathsf{G}} \exists z_{/Y}\psi[V_{x:\,f}]$.
Conversely, if $\mathfrak{A} \models^+_{\mathsf{G}} \exists z_{/Y}\psi[V_{x:\,f}]$, then there is a Y-independent $g : V_{x:\,f} \to A$ such that $\mathfrak{A} \models^+_{\mathsf{G}} \psi[(V_{x:\,f})_{z:\,g}]$. Define $g^* : V \to A$ by $g^*(v) = g(v(x\colon f(v)))$ for all $v \in V$. Then g^* is independent of Y, because g is independent of Y and f is independent of $Z \supseteq Y$. Because x and z are distinct, we have: $(V_{x:\,f})_{z:\,g} = (V_{z:\,g^*})_{x:\,f^*}$. Therefore: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[(V_{z:\,g^*})_{x:\,f^*}]$, and by induction: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_{z:\,g^*}]$ (because f^* is independent of $Z \cup \{z\}$), so (by Y-independence of g^*): $\mathfrak{A} \models^+_{\mathsf{G}} \exists z_{/Y}\psi[V]$.

(-) The following are equivalent:

* $\mathfrak{A} \models_{\mathsf{G}}^{-} \exists z_{/Y}\psi[V_{x:\, f}]$
* $\mathfrak{A} \models_{\mathsf{G}}^{-} \psi[(V_{x:\, f})_{z:\, A}]$
* $\mathfrak{A} \models_{\mathsf{G}}^{-} \psi[(V_{z:\, A})_{x:\, f^*}]$ (note that $(V_{x:\, f})_{z:\, A} = (V_{z:\, A})_{x:\, f^*}$ because x and z are distinct, which follows from the assumption that x does not occur in φ.)
* $\mathfrak{A} \models_{\mathsf{G}}^{-} \psi[V_{z:\, A}]$ by induction hypothesis (because f^* is independent of $Z \cup \{z\}$, the case that z occurs free under the slashes of ψ is covered)
* $\mathfrak{A} \models_{\mathsf{G}}^{-} \exists z_{/Y}\psi[V]$

$\triangleleft$

Corollary 6.2.7 (Adding mute quantifiers) *For any IFG-formula φ: let $\mathfrak{A}$ be a suitable model, $V \subseteq A^X$ a set of valuations for φ in $\mathfrak{A}$. If x is a variable that does not occur in φ nor in X, and Z the set of all free variables of φ that occur under the slashes in φ, then:*

$$\mathfrak{A} \models_{\mathsf{G}}^{\pm} \varphi[V] \textit{ iff } \mathfrak{A} \models_{\mathsf{G}}^{\pm} \exists x_{/Z}\varphi[V] \textit{ iff } \mathfrak{A} \models_{\mathsf{G}}^{\pm} \forall x_{/Z}\varphi[V]$$

Proof: It follows from lemma 6.2.2 that $\mathfrak{A} \models_{\mathsf{G}}^{\pm} \varphi[V]$ iff $\mathfrak{A} \models_{\mathsf{G}}^{\pm} \varphi[V_{x:\, A}]$. Together with lemma 6.2.6 this proves the corollary. $\triangleleft$

Note that for sentences, the set of variables Z is empty, so we may add empty unslashed quantifications ($\exists x, \forall x$, where x as in the corollary).

6.3 Equivalence and Substitution

The definition of the semantics and the results of the previous section all concern the evaluation of an IFG-formula in a given model with a given set of valuations. In this section we move up to a higher level and deal with the question: when are two IFG-formulas *equivalent*? For example, one would be inclined to interpret the last result of the previous section (corollary 6.2.7) as: φ is *equivalent* to $\exists x_{/Z}\varphi$, if we take x and Z as required. We will come back to this.

In [CK99], a natural notion of equivalence for IFG-formulas, *G-equivalence*, is defined as follows (where the 'G' stands for 'Game'):

Definition 6.3.1 (G-equivalence, [CK99, p. 24]) *Two IFG-formulas φ and ψ are* **G-equivalent**, *notation: $\varphi \equiv_G \psi$, if and only if for any suitable model $\mathfrak{A}$ and any set of valuations $V \subseteq A^X$ with domain $X \supseteq Fv(\varphi) \cup Fv(\psi)$ we have both*

$$\mathfrak{A} \models_{\mathsf{G}}^{+} \varphi[V] \Leftrightarrow \mathfrak{A} \models_{\mathsf{G}}^{+} \psi[V], \textit{ and } \mathfrak{A} \models_{\mathsf{G}}^{-} \varphi[V] \Leftrightarrow \mathfrak{A} \models_{\mathsf{G}}^{-} \psi[V].$$

We would like to remark that this notion of equivalence is stronger than the notion of equivalence used by Hintikka (in [Hin96] and other work), where it is defined as having the same positive evaluation only: see the discussion of truth equivalence versus strong equivalence in section 3.3 of this thesis. (We note that Hodges also uses the word equivalence in a weak sense on page 57 of [Hod97c]: he calls two formulas *logically equivalent* if they have the same trumps. To make this equivalence strong, one would also have to require that the formulas have the same *co-trumps*.)

G-equivalence allows us to formally express that double negation cancels (we can now say: $\neg\neg\varphi \equiv_G \varphi$), and this immediately gives us De Morgan's laws (e.g. $\neg\exists x_{/Y}\psi \equiv_G \forall x_{/Y}\neg\psi$) by the fact that universal quantification and conjunction are defined in terms of their duals and negation. These facts will be used in some of our proofs.

One of the elementary properties of first order logic is the fact that the names of the bound variables are not instrumental for the truth or falsity of a formula: e.g. if ψ is a first order formula in which the variable z does not occur, then $\exists x\psi(x) \equiv \exists z\psi(z)$. One would expect this to hold for IFG-formulas as well, and indeed Lemma 3.1(a) of [CK99, p. 26] claims such a result for IFG-formulas, using the straightforward notion of G-equivalence in order to formalize this.

Unfortunately, it turns out that the claim is invalidated by two types of counterexamples: the 'new' variable may introduce or disturb signaling possibilities. This can occur in those situations where the formula and its renamed version are evaluated with valuations whose domain contain the old or the new variable.

The first example will show that the natural notion of G-equivalence is stricter than expected. In fact it invalidates the following claim:

Quote 6.3.2 ([CK99], Lemma 3.1(a)) *For every IFG-formula $\exists x_{/Y}\varphi(x)$ and any variable z that does not occur in $\exists x_{/Y}\varphi(x)$, the following holds:*

$$\exists x_{/Y}\varphi(x) \equiv_G \exists z_{/Y}\varphi(z).$$

The following example demonstrates one problem with this claim:

Example 6.3.3: (Outside signals can be blocked) Let $\varphi(x)$ be the formula $\exists y_{/u}[t = x \wedge u = y]$. Note that the variable z doesn't occur in φ, so we let $\varphi(z)$ be the formula $\exists y_{/u}[t = z \wedge u = y]$. According to the claim quoted above, the formulas $\exists x\varphi(x)$ and $\exists z\varphi(z)$ should be equivalent:

$$\exists x\exists y_{/u}[t = x \wedge u = y] \overset{?}{\equiv}_G \exists z\exists y_{/u}[t = z \wedge u = y]$$

This would mean that their evaluation is the same for all models $\mathfrak{A}$ and sets of valuations V whose domains contain the free variables u and t.

However, let $\mathfrak{A}$ be a model with domain $A = \{0, 1\}$ and let V be the set of valuations $\{utz\colon 000, 010, 101, 111\}$. This situation could occur e.g. in the games for the sentences:

$$\forall u\forall t\exists z[u = z \wedge \exists x\exists y_{/u}[t = x \wedge u = y]] \tag{6.1}$$

and

$$\forall u \forall t \exists z[u = z \wedge \exists z \exists y_{/u}[t = z \wedge u = y]] \tag{6.2}$$

respectively. We note that in the second sentence, the quantification $\exists z$ of $\varphi(z)$ variable z occurs nested in the scope of a previous quantification $\exists z$. We will re-encounter this several times in the rest of this chapter.

We will now verify that $\mathfrak{A} \models^+_{\mathsf{G}} \exists x\varphi(x)[V]$, but $\mathfrak{A} \not\models^+_{\mathsf{G}} \exists z\varphi(z)[V]$:

- $\mathfrak{A} \models^+_{\mathsf{G}} \exists x\varphi(x)[V]$: Eloïse will obviously use the function $f : V \to A$ with $f(v) = v(t)$ to choose a value for x. Continuing the game with

$$V_{x:\,f} = \{utzx\colon 0000, 0101, 1010, 1111\},$$

  she might use $g : V_{x:\,f} \to A$ with $g(w) = w(u)(= w(z))$ to choose a value for y. As there are no two valuations in $V_{x:\,f}$ that coincide out of $\{u\}$, any function and in particular g is independent of $\{u\}$! The functions f, g can easily be seen to be a winning strategy for Eloïse in $\mathsf{G}(\mathfrak{A}, \exists x\varphi(x), V)$.

- $\mathfrak{A} \not\models^+_{\mathsf{G}} \exists z\varphi(z)[V]$: Again, Eloïse has no choice but to use the function $f : V \to A$ with $f(v) = v(t)$ in order to satisfy the atomic formula $(t = z)$. Because z was already in the domain of V, we have to overwrite the old value of z and continue the game with

$$V_{z:\,f} = \{utz\colon 000, 011, 100, 111\}.$$

  Now, when choosing a value for y, the original value of z is no longer available to signal the value of u. Any function $g : V_{z:\,f} \to A$ that is independent of u has to assign the same value to the valuations $(utz\colon 000)$ and $(utz\colon 100)$, and therefore cannot guarantee that $g(v) = v(u)$. So $\mathfrak{A} \not\models^+_{\mathsf{G}} \exists z\varphi(z)[V]$.

◇

This example is a counterexample to the Lemma 3.1(a) of [CK99] quoted above. The example shows that it is sometimes possible to signal with the values of variables that don't occur in the formula (z did not occur in $\exists x\varphi(x)$), but do occur in the domain of the set of valuations X. This signaling opportunity can disappear if we rename a bound variable by an element of X that did not occur in the formula before.

It follows that we cannot rename bound variables just with any variable that does not occur in our formula. We have to make sure that this new variable does not occur in the domain X of the set of valuations V 'under consideration'. But this shows the notion of G-equivalence of definition 6.3.1 to be more tricky than it might have looked at first glance: it contains a quantification over *all* sets of variables V, hence implicitly over all possible domains X containing the free variables of the formulas. With the quantification over *all* such V, the notion of equivalence has become too strict, at least to allow for renaming of bound variables. The example above indicates the general problem: whatever 'new' variable we choose, there is always a set of valuations whose domain contains it. A similar problem holds for a formulation of corollary 6.2.7 in terms of G-equivalence: the variable x in that result is to be chosen *given* a certain domain X of variables.

To clear the way for renaming of bound variables, we define an alternative, weaker notion of equivalence, by restricting the quantification over sets of valuations in the definition of Game equivalence (definition 6.3.1). We will declare two formulas φ and ψ equivalent, if they evaluate identically only with respect to sets of valuations whose domains are in a certain sense *safe* for φ and ψ:

Definition 6.3.4 (Safe sets of valuations) *Let φ be an IFG-formula. A set of valuations $V \subseteq A^X$ is* **safe** *for φ if $Fv(\varphi) \subseteq X$ and $Bv(\varphi) \cap X = \emptyset$.*

Safeness is a condition on the domains X of the sets of valuations V in terms of the free and bound variables of the formula that is evaluated. We would like to loosen the notion of equivalence by only taking safe sets of valuations into account. This means: if we compare two formulas φ, ψ, we restrict the comparison to sets of valuations $V \subseteq A^X$ that are safe for *both* φ and ψ, implying that $Fv(\varphi) \cup Fv(\psi) \subseteq X$ and $(Bv(\varphi) \cup Bv(\psi)) \cap X = \emptyset$. But suppose for example that a free variable of φ occurs bound in ψ, then no sets of valuations can be safe for both formulas! A naive restriction to sets of valuations that are safe for both formulas would then make such formulas trivially equivalent, on the basis of the rather arbitrary choice of variables. For example, it would make $\varphi := (z = z)$ equivalent to $\psi := \forall z[z \neq z]$, which is clearly not what we would intend.

We will define our new notion of 'S-equivalence' (with the 'S' for *safe*) only for pairs of formulas for which safe sets of valuations exist. We call these pairs S-compatible. Note that $Fv(\varphi) \cup Fv(\psi)$ can be abbreviated by $Fv(\varphi \wedge \psi)$, and $Bv(\varphi) \cup Bv(\psi)$ by $Bv(\varphi \wedge \psi)$.

Definition 6.3.5 (S-compatibility) *Two IFG-formulas φ, ψ are called* **S-compatible** *if they satisfy the condition that $Fv(\varphi \wedge \psi) \cap Bv(\varphi \wedge \psi) = \emptyset$.*

The following abbreviation will be useful. We will define 'full' *regularity* later, in definition 6.3.15, but semi-regularity plays a role in the domain of the notion of S-equivalence we define here.

Definition 6.3.6 (semi-regularity) *An IFG-formula φ will be called* **semi-regular** *if no variable occurs both free and bound in φ ($Fv(\varphi) \cap Bv(\varphi) = \emptyset$).*

Note that an IFG-formula is S-compatible with itself precisely if it is semi-regular.

Definition 6.3.7 (S-equivalence) *Two S-compatible IFG-formulas φ, ψ are* **S-equivalent**, *notation: $\varphi \equiv_S \psi$, if and only if for any suitable model $\mathfrak{A}$ and any* **safe** *set of valuations $V \subseteq A^X$ for φ and ψ, we have both*

$$\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V] \Leftrightarrow \mathfrak{A} \models^+_{\mathsf{G}} \psi[V], \text{ and } \mathfrak{A} \models^-_{\mathsf{G}} \varphi[V] \Leftrightarrow \mathfrak{A} \models^-_{\mathsf{G}} \psi[V].$$

S-equivalence is only defined for pairs of formulas that are S-compatible. By this choice, we avoid the situation that S-incompatible pairs of formulas would be trivially S-equivalent, for no sets of valuations V satisfy the precondition in the definition. As a consequence however, S-equivalence is not total, and in particular: $\varphi \equiv_S \varphi$ holds only for IFG-formulas φ that are semi-regular. So, we only have restricted reflexivity. It is easy to verify that S-equivalence is symmetric. We now prove that we also have (restricted) transitivity for S-equivalence, viz. for S-compatible pairs of formulas:

Lemma 6.3.8 (Transitivity of S-equivalence) *Let φ, ψ and χ be IFG-formulas such that all three pairs $\{\varphi, \psi\}$, $\{\psi, \chi\}$ and $\{\chi, \varphi\}$ are S-compatible. Then: if $\varphi \equiv_S \psi$ and $\psi \equiv_S \chi$, then also $\varphi \equiv_S \chi$.*

Proof: Let $\mathfrak{A}$ be a suitable model (for all three formulas). We will first show that any $V \subseteq A^X$ that is safe for φ and χ (i.e. $Fv(\varphi) \cup Fv(\chi) \subseteq X = dom(V)$, and $(Bv(\varphi) \cup Bv(\chi)) \cap X = \emptyset$) can be transformed into a set $V^* \subseteq A^{X^*}$ such that:

1. V^* is safe for φ, χ *and* ψ (i.e. $Fv(\varphi) \cup Fv(\chi) \cup Fv(\psi) \subseteq X^* = dom(V^*)$ and $(Bv(\varphi) \cup Bv(\chi) \cup Bv(\psi)) \cap X^* = \emptyset$);
2. $\mathfrak{A} \models^{\pm}_{G} \varphi[V]$ iff $\mathfrak{A} \models^{\pm}_{G} \varphi[V^*]$
3. $\mathfrak{A} \models^{\pm}_{G} \chi[V]$ iff $\mathfrak{A} \models^{\pm}_{G} \chi[V^*]$

First step is to make sure that any bound variable of ψ that occurs in X, no longer occurs in X^*. So, suppose there is a variable $z \in X$ that occurs bound in ψ. By safeness of V, we know that z does not occur *bound* in either φ or χ. Also, by S-compatibility of φ, ψ, and of ψ, χ respectively, we know that the bound variables of ψ do not occur free in φ or χ either. So: z does not occur in φ nor χ. In that case, we can safely 'rename' the variable z in the domain X of V by a variable z' that does not occur in any of the formulas φ, ψ or χ: let $X' := (X - \{z\}) \cup \{z'\}$, and make $V' \subseteq A^{X'}$ the set of all v' defined from some $v \in V$ by $v'(x) := v(x)$ for all $x \in X - \{z\}$ and $v'(z') := v(z)$. Then V' is still safe for φ, χ and we claim:

$$\mathfrak{A} \models^{\pm}_{G} \varphi[V'] \text{ iff } \mathfrak{A} \models^{\pm}_{G} \varphi[V']$$

$$\mathfrak{A} \models^{\pm}_{G} \chi[V'] \text{ iff } \mathfrak{A} \models^{\pm}_{G} \chi[V']$$

This is an elementary insight: if there happens to be some kind of signaling possible through values assigned to a variable that does not occur in the formulas, the *name* of that variable is irrelevant. We don't think the formal proof contributes to the understanding of it, so we omit it here. We repeat this operation until all bound variables of ψ that occurred in $X = dom(V)$ are renamed. Let's call the result V'' and its domain X''.

Second step is to make sure that $X^* = dom(V^*)$ will contain all free variables of ψ. Let $Z := Fv(\psi) - X''$ be the set of free variables of ψ that are not yet in the domain X'' of V''. Define $X^* := X'' \cup Z$ and let $V^* := V'' \times A^Z$. Then V^* is safe for ψ, and because the free variables of ψ do not occur bound in φ or χ (by S-compatibility of both formulas with ψ): V^* is also safe for φ and ψ. Furthermore, by lemma 6.2.2:

$$\mathfrak{A} \models^{\pm}_{G} \varphi[V''] \text{ iff } \mathfrak{A} \models^{\pm}_{G} \varphi[V^*]$$

$$\mathfrak{A} \models^{\pm}_{G} \chi[V''] \text{ iff } \mathfrak{A} \models^{\pm}_{G} \chi[V^*]$$

So, for every V that is safe for φ and χ, we can construct a V^* that satisfies the requirements 1-3 above.

Now suppose that $\varphi \equiv_S \psi$ and $\psi \equiv_S \chi$, and that $V \subseteq A^X$ is safe for φ, χ. Then the following are equivalent:

- $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V]$
- $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi[V^*]$ (by construction)
- $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \psi[V^*]$ (by $\varphi \equiv_S \psi$)
- $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \chi[V^*]$ (by $\psi \equiv_S \chi$)
- $\mathfrak{A} \models^{\pm}_{\mathsf{G}} \chi[V]$ (by construction)

Hence: $\varphi \equiv_S \chi$. $\triangleleft$

Despite the restrictions on reflexivity and transitivity, we call the relation $\equiv_S$ 'S-*equivalence*'. We made this choice, because apart from its mathematical connotations, the word *equi-valence*, 'the same values', expresses best what the relation means (with the subscript 'S' indicating the restriction to *safe* models).

Note that for S-compatible pairs of formulas, G-equivalence *implies* S-equivalence. In particular, all pairs of IFG-*sentences* are naturally S-compatible, and due to corollary 6.2.3, we have

Lemma 6.3.9 *For all IFG-sentences φ, ψ: $\varphi \equiv_G \psi$ if and only if $\varphi \equiv_S \psi$.*

In Example 6.3.3, which proved that G-equivalence is too strong to allow for renaming of bound variables, the fact that the 'new' bound variable z occurred in the domain X of V was crucial. This situation will no longer occur if we only consider safe V, i.e. if we compare (S-compatible) formulas by S-equivalence.

Also, we can now indeed state corollary 6.2.7 in terms of S-equivalence:

Corollary 6.3.10 (Addition of mute quantifiers reformulated) *For every semi-regular IFG-formula φ: if x is a variable that does not occur in φ, and Z is the set of all free variables of φ that occur under the slashes in φ, then*

$$\varphi \equiv_S \exists x_{/Z}\varphi \text{ and } \varphi \equiv_S \forall x_{/Z}\varphi$$

The introduction of S-equivalence solved the first problem with renaming of bound variables, but there is a second obstacle we have to deal with. The situation of example 6.3.3 was related to the occurrence of a quantification Qx within the scope of another quantification $Q'x$, as can be recognized in the sentence (6.2) in the example. We call this *nested* quantification over x. The following variation of our previous example shows how this can be a problem by itself:

Example 6.3.11: (Inside signals can be unblocked) Let φ be the formula $\exists x \exists y_{/u}[t = x \wedge u = y]$ (this is the formula $\exists x\varphi(x)$ from Example 6.3.3). Now consider the formula $\exists x\varphi$. Suppose we want to replace the first (vacuous) quantification over x in $\exists x\varphi$ by a quantification over z to get $\exists z\varphi$. These formulas are S-compatible, hence we can ask whether they are S-equivalent:

$$\exists x\varphi \overset{?}{\equiv}_S \exists z\varphi$$

Let $\mathfrak{A}$ be any model with domain $A = \{0,1\}$, and let V be the set of valuations $\{ut\colon 01, 11\}$. This situation could occur in games on the sentences $\exists t \forall u \exists x \varphi$ and $\exists t \forall u \exists z \varphi$. Note that V is safe for $\exists x \varphi$ and $\exists z \varphi$.

Now consider first $\exists z \varphi$. Informally, by choosing the value of u for z (continue with $\{utz\colon 010, 111\}$), the value of t for x (continue with $\{utxz\colon 0110, 1111\}$) and the value of z for y respectively, Eloïse has a winning strategy: the resulting set of valuations $\{utxyz\colon 01100, 11111\}$ satisfies the propositional part $[t = x \wedge u = y]$. In this case, she can use the empty quantification over z to signal the value of u to y. So, $\mathfrak{A} \models^{+}_{\mathsf{G}} \exists z \varphi[V]$.

In the case of the formula $\exists x \varphi$, Eloïse's best choice is to play the same strategy as in the previous case. Again, Eloïse can try to signal the value of u by the first, empty quantification $\exists x$ (and continue with $\{utx\colon 010, 111\}$). But at the second quantification $\exists x$, a *new* value is chosen for x. In order to satisfy the atomic formula $t = x$, Eloïse has to continue with $V' := \{utx\colon 011, 111\}$, thereby overwriting the attempted signal. As the last choice has to be independent of u, and both valuations in V' coincide out of u, a u-independent function f has to assign the same value to both valuations. This means that one valuation in the resulting set $V'_{y\colon f}$ will not satisfy the atomic formula $u = y$. So, $\mathfrak{A} \not\models^{+}_{\mathsf{G}} \exists x \varphi[V]$.

The renaming of the (first occurrence of the) variable x into the new variable z has changed the formula $\exists x \varphi$ from not-true into true, even with respect to a safe set of valuations. ◇

Whereas in Example 6.3.3 the new variable z blocked signaling, we see in Example 6.3.11 that it can also allow for new signaling possibilities. The first example asked for conditions on the implicit domains X of the sets of valuations V in the definition of equivalence, the latter can be solved by a syntactical condition on the formulas involved, viz. an extra condition on the variable that is to be replaced. These considerations lead to the following theorem, which replaces Lemma 3.1(a) of [CK99, p. 26]:

Lemma 6.3.12 (Renaming of Bound Variables) *If $\varphi(x)$ is an IFG-formula in which the variable x does not occur bound, then for every variable z that does not occur in $\exists x_{/Y} \varphi(x)$:*

$$\exists x_{/Y} \varphi(x) \equiv_S \exists z_{/Y} \varphi(z)$$

Proof: Note that $\exists x_{/Y} \varphi(x)$ and $\exists z_{/Y} \varphi(z)$ are S-compatible because z does not occur free in $\exists x_{/Y} \varphi(x)$. Let $\mathfrak{A}$ be a suitable model for φ and $V \subseteq A^X$ be a set of valuations that is safe for $\exists x_{/Y} \varphi(x)$ and $\exists z_{/Y} \varphi(z)$. Then:

(+) $\mathfrak{A} \models^{+}_{\mathsf{G}} \exists x_{/Y} \varphi(x)[V]$ iff $\mathfrak{A} \models^{+}_{\mathsf{G}} \varphi(x)[V_{x\colon f}]$ for some $f : V \to A$ independent of Y. Because x does not occur bound in $\varphi(x)$ and because z does not occur in $\varphi(x)$ nor in X, we have for any such f: $\mathfrak{A} \models^{+}_{\mathsf{G}} \varphi(x)[V_{x\colon f}]$ iff $\mathfrak{A} \models^{+}_{\mathsf{G}} \varphi(z)[V_{z\colon f}]$. The last clause is equivalent to $\mathfrak{A} \models^{+}_{\mathsf{G}} \exists z_{/Y} \varphi(z)[V]$.

(-) $\mathfrak{A} \models^{-}_{\mathsf{G}} \exists x_{/Y} \varphi(x)[V]$ iff $\mathfrak{A} \models^{-}_{\mathsf{G}} \varphi(x)[V_{x\colon A}]$ iff $\mathfrak{A} \models^{-}_{\mathsf{G}} \varphi(z)[V_{z\colon A}]$ (because neither x nor z are in X) iff $\mathfrak{A} \models^{-}_{\mathsf{G}} \exists z_{/Y} \varphi(z)[V]$

◁

Technically, this proves only that variables that are bound by a formula-initial quantification may be renamed. In order to generalize this to arbitrary bound variables (occurring deeper in the formula), we need a substitution principle. In fact, the substitution principle Theorem 2.2 of [CK99, p. 24] for G-equivalent formulas – "substituting G-equivalent formulas in a given IFG-formula yields G-equivalent formulas" – is correct. But G-equivalence was too strict for renaming of bound variables, so we are looking for a similar substitution principle for S-equivalence: if φ occurs as subformula of θ, and if φ is equivalent to ψ, under which conditions is the result $\theta[\psi/\varphi]$ of replacing the occurrence of φ in θ by ψ S-equivalent to θ?

The following example, very similar to our previous one, shows that substitution for S-equivalence is not trivial:

Example 6.3.13: (Substitution may block signals) Elaborating on Example 6.3.11, let $\theta := \exists z\varphi$ where φ is again the formula $\exists x\exists y_{/u}[t = x \wedge u = y]$. Let $\psi := \exists z\exists y_{/u}[t = z \wedge u = y]$. Then $\varphi \equiv_S \psi$ by lemma 6.3.12.

Now consider $\theta[\psi/\varphi] = \exists z\psi = \exists z\exists z\exists y_{/u}[t = z \wedge u = y]$. By taking any model $\mathfrak{A}$ with domain $\{0, 1\}$, and $V := \{ut\colon 01, 11\}$, we see that $\mathfrak{A} \models^+_{\mathsf{G}} \exists z\varphi$ (because z can be used to signal the value of u to y), but $\mathfrak{A} \not\models^+_{\mathsf{G}} \exists z\psi$ (because the quantification over z in ψ blocks the signal).

So: $\theta \not\equiv_S \theta[\psi/\varphi]$. ◇

In the result of the substitution, a signaling possibility is blocked because the quantification over z in ψ comes to fall within the scope of another quantification over z. We also have another problem: the result of a substitution may be S-incompatible with the original:

Example 6.3.14: (S-incompatibility by substitution) Let φ and ψ be as in the previous example, and let θ' be the formula $(z = z) \wedge \varphi$. Then $\theta'[\psi/\varphi] := (z = z) \wedge \psi$. Because in the latter formula the variable z occurs both free and bound, the pair $\{\theta', \theta'[\psi/\varphi]\}$ is not S-compatible. ◇

The S-incompatibility in the example is more a problem of the formula $\theta'[\psi/\varphi]$ by itself than of the *pair* $\theta', \theta'[\psi/\varphi]$.

These two examples motivate the following definition:

Definition 6.3.15 (Regularity) *An IFG-formula φ is called* **regular** *if*

- *no variable occurs both free and bound in φ (i.e. φ is semi-regular),* and
- *within the scope of a quantification there is no other quantification over the* same *variable.*

Let's make some clarifying remarks on this definition.

The first clause in the definition ensures that safe sets of valuations exist for regular formulas, and the second clause ensures that if we start evaluating a regular IFG-formula with a safe set of valuations V, then all sets of valuations in the inductive evaluation steps are safe for the corresponding subformulas. This will be needed for the safe application of our induction hypotheses in some of our proofs.

Regular formulas are the formulas that can safely be evaluated inductively using safe sets of valuations.

Note that the second clause does not exclude formulas in which *disjoint* quantification over the same variable occurs, as for example: $\forall x P(x) \vee \neg\forall x P(x)$. These formulas don't cause the type of problems we have pointed out in our examples: only with *nested* quantification are earlier assignments to a variable overwritten, thereby possibly blocking signals. Furthermore, the second clause not only implies that the first clause holds for any of the subformulas, but it also excludes nested quantification if one of the quantifiers is empty, like in the formula $\theta[\psi/\varphi] = \exists z \exists z \exists y_{/u}[t = z \wedge u = y]$ of Example 6.3.13. We exclude this type of situation for the reason mentioned above: we want to guarantee for our induction arguments, that all sets of valuations encountered in the inductive evaluation of a regular formula with respect to a safe set of valuations, are safe for the corresponding subformulas. This would not be the case in the second step of the evaluation of the formula $\theta[\psi/\varphi]$: we then evaluate $\psi := \exists z \exists y_{/u}[t = z \wedge u = y]$ with a z-variation V_z of $V \subseteq A^X$. $V_z \subseteq A^{X \cup \{z\}}$ has the bound variable z of ψ in its domain, hence is not safe for ψ even if V was safe for $\theta[\psi/\varphi]$.

It follows easily from the definition of regularity, that any subformula of a regular formula is regular. For sentences, regularity reduces to just the second clause of this definition. Regularity of a single formula can be seen in relation to S-compatibility of pairs in the following manner:

- By the first clause of the definition, regularity of an IFG-formula φ implies that the pair $\{\varphi, \varphi\}$ is S-compatible (hence: reflexivity of S-equivalence holds for regular formulas).

- If $\varphi \wedge \psi$ is a regular formula, then so are φ and ψ, and the pair $\{\varphi, \psi\}$ is then S-compatible. Note that regularity of $\varphi \wedge \psi$ is stronger than S-compatibility of the pair $\{\varphi, \psi\}$ by the second clause of the definition of regularity: this requires that neither φ nor ψ contain nested quantification over the same variable, which is no requirement for S-compatibility of $\{\varphi, \psi\}$.

The notion of regularity enables us to state the following substitution principle for S-equivalence:

Theorem 6.3.16 (Substitution of S-equivalents) *For every regular IFG-formula θ, and S-compatible formulas φ, ψ: if the result $\theta[\psi/\varphi]$ of replacing an occurrence of φ in θ by ψ is regular, then*

$$\varphi \equiv_S \psi \textit{ implies } \theta \equiv_S \theta[\psi/\varphi]$$

Proof: Let φ, ψ be S-compatible formulas such that $\varphi \equiv_S \psi$. Let $\theta[\psi/\varphi]$ denote (ambiguously) the result of substituting either one, several, or no occurrences of φ in ψ.

First, we ensure in general that from the regularity assumptions on θ and $\theta[\psi/\varphi]$, and the fact that φ, ψ are S-compatible, it follows that $\theta, \theta[\psi/\varphi]$ are S-compatible: In case of no substitutions, this follows by regularity of θ. Otherwise, suppose they are not S-compatible. Then there would be a variable x in the intersection of $Fv(\theta \wedge \theta[\psi/\varphi])$ and $Bv(\theta \wedge \theta[\psi/\varphi])$. Because of the regularity assumptions on θ and $\theta[\psi/\varphi]$, it then follows that this can only be the case if either $x \in Fv(\varphi) \cap Bv(\psi)$ or $x \in Fv(\psi) \cap Bv(\varphi)$. Both are impossible by the assumption that φ, ψ are S-compatible. So, θ and $\theta[\psi/\varphi]$ are S-compatible.

We now prove the claim by induction on the structure of θ. For regular θ in which φ does not occur, we have $\theta[\psi/\varphi] = \theta$ and $\theta \equiv_S \theta$ by regularity of θ. For regular θ in which φ occurs (this implies regularity of φ), start the induction with $\theta = \varphi$. Then $\theta[\psi/\varphi] = \psi$ or $\theta[\psi/\varphi] = \varphi$ (in case of no substitutions). In the first case, we have $\theta \equiv_S \theta[\psi/\varphi]$ by assumption, in the second case it follows from regularity of φ.

Now assume that θ is regular, that φ is a (real) subformula of θ, and that $\theta[\psi/\varphi]$ is regular.

For the induction, assume that for all regular formulas θ' of complexity smaller than the complexity of θ we have: $\theta' \equiv_S \theta'[\psi/\varphi]$ if $\theta'[\psi/\varphi]$ is regular. Note that regularity of θ and $\theta[\psi/\varphi]$ imply regularity of all their subformulas. This is used (but not stated explicitly) in the following cases:

($\neg$) $\theta = \neg\theta'$: this case follows immediately from the induction hypothesis.

($\vee$) $\theta = \theta_1 \vee_{/Y} \theta_2$: then $\theta[\psi/\varphi] = \theta_1[\psi/\varphi] \vee_{/Y} \theta_2[\psi/\varphi]$. By induction: $\theta_i \equiv_S \theta_i[\psi/\varphi]$ for both $i = 1, 2$. From this it easily follows that $\theta \equiv_S \theta[\psi/\varphi]$.

($\exists$) $\theta = \exists x_{/Y}\theta'$: then $\theta[\psi/\varphi] = \exists x_{/Y}\theta'[\psi/\varphi]$. Note that x nor the variables in Y occur bound in θ nor in $\theta[\psi/\varphi]$ by regularity. Now, let $\mathfrak{A}$ be a suitable model and V a safe set of valuations for $\theta, \theta[\psi/\varphi]$ (such V exists by S-compatibility of $\theta, \theta[\psi/\varphi]$).

(+) The following are equivalent:

* $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\theta'[V]$
* $\mathfrak{A} \models^+_{\mathsf{G}} \theta'[V_{x:\,f}]$ for some Y-independent $f : V \to A$. (by definition) Note that $V_{x:\,f}$ is safe for both θ' and $\theta'[\psi/\varphi]$ since x occurs bound in neither φ nor ψ by regularity of θ and $\theta[\psi/\varphi]$.
* $\mathfrak{A} \models^+_{\mathsf{G}} \theta'[\psi/\varphi][V_{x:\,f}]$ (by induction hypothesis)
* $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\theta'[\psi/\varphi][V]$.

(-) The following are equivalent:

* $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\theta'[V]$
* $\mathfrak{A} \models^+_{\mathsf{G}} \theta'[V_{x:\,A}]$ (by definition) Note that $V_{x:\,A}$ is again safe for both θ' and $\theta'[\psi/\varphi]$
* $\mathfrak{A} \models^-_{\mathsf{G}} \theta'[\psi/\varphi][V_{x:\,A}]$ (by induction hypothesis)
* $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\theta'[\psi/\varphi][V]$.

This means that $\theta \equiv_S \theta[\psi/\varphi]$.

◁

With this substitution result, we can generalize lemma 6.3.12:

Theorem 6.3.17 (Renaming of bound variables) *If φ' is a regular IFG-formula obtained from a regular IFG-formula φ by renaming of bound variables, then $\varphi \equiv_S \varphi'$.*

Proof: Let φ be a regular IFG-sentence, and let ψ' be the result of application of renaming of bound variables (lemma 6.3.12) to a subformula ψ of φ. Then $\psi \equiv_S \psi'$, and if $\varphi' = \varphi[\psi'/\psi]$ is regular, then by the Substitution Principle 6.3.16: $\varphi \equiv_S \varphi'$. $\lhd$

Corollary 6.3.18 *If φ' is a regular IFG-sentence obtained from a regular IFG-sentence φ by renaming of bound variables, then $\varphi \equiv_G \varphi'$.*

Proof: By the previous theorem and by the fact that G-equivalence and S-equivalence coincide for sentences (lemma 6.3.9). $\lhd$

The substitution principle for S-equivalence is a crucial component of the proof of the prenex normal form that we will prove (theorem 6.5.1). In the next section we will study another important component of this theorem: quantifier extraction.

6.4 Quantifier extraction

Whereas most of our previous examples focused on the interpretation of the slashed quantifiers in IFG, we will now show how similar signaling situations can occur with connectives. The results of this section are all inspired by a detailed study of 'quantifier extraction'. In [CK99], quantifier extraction for IFG-formulas is phrased as follows (with Q^d meaning the dual of the quantifier Q, e.g. $Q^d = \exists$ if $Q = \forall$):

Quote 6.4.1 (Lemma 3.1(b)-(d), [CK99, p. 26]) *For any IFG-formula φ*

(b) $\neg Qx_{/Y}\varphi \equiv_G Q^d x_{/Y}\neg\varphi$

Let $\psi|x$ denote the result of adding to all quantifiers in ψ the independence condition $/x$. Then:

(c) $[Qx_{/Y}\phi \vee \psi] \equiv_G Qx_{/Y}[\phi \vee \psi|x]$

(d) $[Qx_{/Y}\phi \wedge \psi] \equiv_G Qx_{/Y}[\phi \wedge \psi|x]$

Part (b) of this lemma follows easily from our definition of $\forall$ in terms of $\exists$, and the fact that double negation naturally cancels.

In parts (c) and (d), $\psi|x$ is defined to be the formula resulting from making all *quantifiers* in ψ independent of x. Note that we introduced the notation $\psi_{/x}$ to denote the formula resulting from making all quantifiers *and* connectives in ψ independent of x, see definition 6.2.4.

After all examples with signaling of values from quantifiers to quantifiers, one becomes suspicious: could the extracted quantifier $Qx_{/Y}$ under whose scope $\psi|x$

comes to fall, not signal values to the *connectives* in $\psi|x$? And, could the extracted quantifier not signal values to the main connective of the formula $(Qx_{/Y}\varphi \vee \psi)$? After all, by the extraction of the quantifier, the order of the moves in the game changes and thereby the available information at these moves.

Example 6.4.2: (Extracted quantifier may produce inside signal) Consider the following instance of Lemma 3.1(c) of [CK99], where the quantifier $\forall x$ is extracted from

$$\forall x[x \neq z] \vee \exists u_{/z}[u = z \vee_{/z} u \neq z].$$

We take a model $\mathfrak{A}$ with domain $A = \{0, 1\}$ and $V := \{z\colon 0, 1\}$. One can easily verify that

$$\mathfrak{A} \not\models^+_{\mathsf{G}} (\forall x[x \neq z] \vee \exists u_{/z}[u = z \vee_{/z} u \neq z])[V].$$

The left disjunct is not made true by any valuation, so Eloïse has to choose the right disjunct for all valuations in V. For $\exists u_{/z}$, a valid strategy must extend both valuations in V with the assignment of the same $c \in \{0, 1\}$ to u, so that we continue with $V_u = \{zu\colon 0c, 1c\}$. A valid choice for the disjunction $\vee_{/z}$ has to pick the same disjunct for both valuations in V_u, but each disjunct is only satisfied by one of the valuations. So, Eloïse does not have a winning strategy.

After extraction of the quantifier $\forall x$ by the quoted Lemma 3.1(c), we get the formula

$$\forall x([x \neq z] \vee \exists u_{/zx}[u = z \vee_{/z} u \neq z]).$$

Because Abélard now chooses the value for x before Eloïse makes her choices, she now has more information (while he has less). And indeed, the extracted formula can be seen to be true: Eloïse can choose the left disjunct for the valuations in $V_{x\colon A}$ that satisfy $(x \neq z)$ $(V_1 = \{xz\colon 01, 10\})$, and the right disjunct for the rest $(V_2 = \{xz\colon 00, 11\})$. At $\exists u_{/xz}$ she extends both valuations with the assignment of (e.g.) 0 to u: $(V_2)_u = \{xzu\colon 000, 110\}$. At $\vee_{/z}$ she can now signal the value of z by the value of x and choose the left disjunct if $(u = z)$, and right otherwise. Together, these choices are a winning strategy for Eloïse:

$$\mathfrak{A} \models^+_{\mathsf{G}} \forall x([x \neq z] \vee \exists u_{/zx}[u = z \vee_{/z} u \neq z]).$$

If we compare the game on the original formula with the extracted one, there were two points in the latter game where Eloïse profited from the availability of the value assigned to x: at the main disjunction $\vee$, and at the slashed disjunction in the right subformula $\vee_{/z}$.

To be safe, we could solve this new information flow, by slashing *all* disjunctions that come to fall under the scope of the extracted quantifier to get the equivalent formula:

$$\forall x([x \neq z] \vee_{/x} \exists u_{/zx}[u = z \vee_{/zx} u \neq z]).$$

But, in this particular case, adding a slash to just one of them, turns out to give equivalent formulas too:

$$\forall x([x \neq z] \vee \exists u_{/zx}[u = z \vee_{/zx} u \neq z]) \text{ and also:}$$

$$\forall x([x \neq z] \vee_{/x} \exists u_{/zx}[u = z \vee_{/z} u \neq z]).$$

◇

The example above contradicts the quoted lemma 3.1(c)-(d) in [CK99]. But a closer look at the proof given in [CK99] learns that in fact, in $\psi|x$ the connectives are meant to be made independent of x, just like the quantifiers. (This can be recognized in the "Reciprocally"-parts of the proofs of (c) and (d), [CK99, p. 27]: the definition of the $g_\theta(w)$ as $f_{\theta|x}(w, c)$ for arbitrary c is not sound unless we make the connectives independent of x as well.)

So, this suggests that we could restore lemma 3.1(c) from [CK99] by redefining $\psi|x$ as our $\psi_{/x}$. But observe that this result is formulated in terms of G-equivalence. And like in the case of Renaming of bound variables, which we discussed in the previous section, it turns out that this notion of equivalence is too strict:

Example 6.4.3: (Extracted quantifier may block outside signals) If $\mathfrak{A}$ is a model with domain $\{0, 1\}$, then

$$\mathfrak{A} \models^+_{\mathsf{G}} \forall x[x \neq x] \vee \exists y_{/z}[y = z]\ [\{zx\colon 00, 11\}]$$

because Eloïse may use the strategy to play 'always right' at the main disjunction, and use the value of x to signal the value of z, in order to ensure satisfaction of the atomic formula $y = z$.

However: $\mathfrak{A} \not\models^+_{\mathsf{G}} \forall x[(x \neq x) \vee \exists y_{/z}[y = z]]\ [\{zx\colon 00, 11\}]$. To see this, suppose:

$$\mathfrak{A} \models^+_{\mathsf{G}} \forall x[(x \neq x) \vee \exists y_{/z}[y = z]]\ [\{zx\colon 00, 11\}].$$

This would respectively imply

$$\mathfrak{A} \models^+_{\mathsf{G}} (x \neq x) \vee \exists y_{/z}[y = z]\ [\{zx\colon 00, 01, 10, 11\}]$$

and (because no valuation can satisfy the atomic formula $x \neq x$):

$$\mathfrak{A} \models^+_{\mathsf{G}} \exists y_{/z}[y = z]\ [\{zx\colon 00, 01, 10, 11\}],$$

The latter is not the case, because the independence restriction on the choice of y requires the valuations $(zx : 00)$ and $(zx : 10)$ to be extended with the same value of z. These extensions can then not all satisfy the atomic formula $y = z$.

Note that we did not even add the extra independence condition with respect to x to the quantification $\exists y_{/z}$, as would be prescribed by the quoted theorem 3.1(c) of [CK99]. That would be to prevent *new* signaling possibilities. We see here however that old possibilities can also be disturbed. By extracting the quantification $\forall x$ over the disjunction, the signal provided to the right disjunct by the initial value of x is overwritten. ◇

This example shows that G-equivalence is too strict a notion of equivalence for quantifier extraction. Just like in example 6.3.3, it is crucial that the variable x, bound by the extracted quantifier, happened to be in the domain of the set of valuations we started the evaluation with: the set of valuations was not *safe* for the formulas in the example.

Example 6.3.3 gave rise to the weaker alternative of S-equivalence to replace G-equivalence in the original formulation of renaming of bound variables. We will use this solution again in the case of quantifier extraction.

Before we formulate and prove a restored version of Quantifier Extraction for IFG-formulas, we raise two new questions. First, in the original formulation of Quote 6.4.1, the main connective of the formula does not get a slash for x after the extraction. Neither did we find examples where this would be necessary. Why not? And second, can we then generalize the lemma to formulas with a slashed main connective, i.e.

$$Qx_{/Y}\varphi \vee_{/Z} \psi \stackrel{?}{\equiv}_F Qx_{/Y}(\varphi \vee_{/Z} \psi_{/x})$$

$$Qx_{/Y}\varphi \wedge_{/Z} \psi \stackrel{?}{\equiv}_F Qx_{/Y}(\varphi \wedge_{/Z} \psi_{/x})$$

The answer to this latter question may seem to be an easy 'yes', but the following counterexample shows again that we should be careful not to judge too fast:

Example 6.4.4: (Slash at main connective) Let $\mathfrak{A}$ again be a model with domain $\{0,1\}$, and $V = \{z\colon 0,1\}$. Then we have

$$\mathfrak{A} \not\models_{\mathsf{G}}^{+} \exists x[z=0] \vee_{/z} [z \neq 0][V]$$

but if we extract the empty quantifier $\exists x$:

$$\mathfrak{A} \models_{\mathsf{G}}^{+} \exists x(z=0 \vee_{/z} z \neq 0)[V].$$

In the latter case Eloïse wins by choosing the value for x equal to the value of z, followed by signaling the value of z by x at the main connective $\vee_{/z}$. Of course:

$$\mathfrak{A} \not\models_{\mathsf{G}}^{+} \exists x(z=0 \vee_{/zx} z \neq 0)[V].$$

◇

Apparently, if the main disjunction of the formula is slashed, we need to add x under the slash after the extraction of $\exists x_{/Y}$. Our first question, why this is not necessary if the main connective isn't slashed, will be answered after we have proven the following corrected and generalized version of the quantifier extraction lemma 3.1(b)-(d) of [CK99].

Lemma 6.4.5 (Quantifier Extraction over $\vee$) *For any semi-regular IFG-formula of the form $Qx_{/Y}\varphi \vee_{/Z} \psi$, where x does not occur in ψ, we have:*

(c') $(\exists x_{/Y}\varphi \vee_{/Z} \psi) \equiv_S \exists x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})$

(d') $(\forall x_{/Y}\varphi \vee_{/Z} \psi) \equiv_S \forall x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})$

Proof: For both (c') and (d'), assume that the left formula $\theta = Qx_{/Y}\varphi \vee_{/Z} \psi$ is semi-regular, and that x does not occur in ψ. Let $\theta' = Qx_{/Y}(\varphi \vee_{/Zx} \psi_{/x})$ denote the right formula. Because θ' has the same free variables and the same bound variables as θ, it follows immediately that the pair $\{\theta, \theta'\}$ is S-compatible.

From the fact that θ is semi-regular, it already follows that x does not occur *free* in ψ. But we still need to require that x does not occur (bound) in ψ, if only because otherwise the subformula $\psi_{/x}$ of θ' would not be well-defined.

Now let $\mathfrak{A}$ be a suitable model and $V \subseteq A^X$ safe for $Qx_{/Y}\varphi \vee_{/Z} \psi$. Note that this implies that $x \notin X$, which is used in every application of lemma 6.2.5, and when taking projections to X.

(c', +) From left to right: $\mathfrak{A} \models^+_{\mathsf{G}} (\exists x_{/Y}\varphi \vee_{/Z} \psi)[V]$ iff $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\varphi[V_1]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_2]$ for V_i Z-saturated in V and $V_1 \cup V_2 = V$. By lemma 6.2.1, we may take the V_i to be disjoint.

Now let $f : V_1 \to A$ be Y-independent such that $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_1)_{x:\,f}]$. We make a function $g : V_2 \to A$ that is Y-independent, and such that $g(v) = f(w)$ if $v \sim_Y w$ for $v \in V_2$ and $w \in V_1$. Lemma 6.2.5 implies then that $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_2)_{x:\,g}]$. Let $h := f \cup g : V \to A$, then h is Y-independent, $(V_1)_{x:\,h} = (V_1)_{x:\,f}$, $(V_2)_{x:\,h} = (V_1)_{x:\,g}$, and $V_{x:\,h} = (V_1)_{x:\,f} \cup (V_2)_{x:\,g}$.

To prove that the $(V_i)_{x:\,h}$ are Zx-saturated in $V_{x:\,h}$: suppose $v(x\colon h(v))$ coincides with $w(x\colon h(w))$ out of Zx, and that $v(x\colon h(v)) \in (V_i)_{x:\,h}$. Then $v \in V_i$ and $v \sim_Z w$, hence, by Z-saturation of V_i: $w \in V_i$, so: $w(x\colon h(w)) \in (V_i)_{x:\,h}$. From this we may conclude $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,h}]$, and therefore: $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})[V]$

Conversely: assume $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})[V]$, and let $f : V \to A$ be Y-independent such that $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,f}]$. Let $(V_{x:\,f})_1, (V_{x:\,f})_2$ be Zx-saturated subsets of $V_{x:\,f}$, such that $V_{x:\,f} = (V_{x:\,f})_1 \cup (V_{x:\,f})_2$ and $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_{x:\,f})_1]$, $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_{x:\,f})_2]$. Now define $V_i := \{v \in V | v(x/f(v)) \in (V_{x:\,f})_i\}$ for $i = 1, 2$ (V_i contains the projections of the elements of $(V_{x:\,f})_i$ to X) It is easily verified that $V_1 \cup V_2 = V$ and that $(V_{x:\,f})_i = (V_i)_{x:\,f}$. Hence: $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_1)_{x:\,f}]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_2)_{x:\,f}]$. So: $\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}\varphi[V_1]$ (because f is Y-independent), and: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_2]$ (by lemma 6.2.5).

To prove that the V_i are Z-saturated in V: if $v \sim_Z w$ for some $v \in V_i, w \in V$, then $v(x\colon f(v)) \in (V_{x:\,f})_i$ and $v(x\colon f(v)) \sim_{Zx} w(x\colon f(w))$. By Zx-saturation of $(V_{x:\,f})_i$, it follows that $w(x\colon f(w)) \in (V_{x:\,f})_i$, and hence $w \in V_i$. We may now conclude: $\mathfrak{A} \models^+_{\mathsf{G}} (\exists x_{/Y}\varphi \vee_{/Z} \psi)[V]$

(c', −) $\mathfrak{A} \models^-_{\mathsf{G}} (\exists x_{/Y}\varphi \vee_{/Z} \psi)[V]$ iff $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}\varphi[V]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V]$, iff $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[V_{x:\,A}]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \psi_{/x}[V_{x:\,A}]$ (by lemma 6.2.5). This is equivalent to: $\mathfrak{A} \models^-_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,A}]$, and equivalently: $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})[V]$.

(d', +) $\mathfrak{A} \models^+_{\mathsf{G}} (\forall x_{/Y}\varphi \vee_{/Z} \psi)[V]$, iff $\mathfrak{A} \models^+_{\mathsf{G}} \forall x_{/Y}\varphi[V_1]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_2]$ with V_i Z-saturated in V and $V_1 \cup V_2 = V$. Equivalently: $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_1)_{x:\,A}]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_2)_{x:\,A}]$ (by lemma 6.2.5). The sets $(V_i)_{x:\,A}$ are Z-saturated in $V_{x:\,A}$, hence Zx-saturated, and $(V_1)_{x:\,A} \cup (V_2)_{x:\,A} = V_{x:\,A}$, so: $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,A}]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \forall x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})[V]$.

Conversely, suppose $\mathfrak{A} \models^+_{\mathsf{G}} \forall x_{/Y}(\varphi \vee_{/Zx} \psi)[V]$, so that $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,A}]$. Then $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_{x:\,A})_1]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_{x:\,A})_2]$ for Zx-saturated subsets $(V_{x:\,A})_i$ of $V_{x:\,A}$ with $V_{x:\,A} = (V_{x:\,A})_1 \cup (V_{x:\,A})_2$.

Let V_i be the projection of $(V_{x:\,A})_i$ to X. By Zx-saturation of $(V_{x:\,A})_i$, we know that any x-variant of $v \in (V_{x:\,A})_i$ is also in $(V_{x:\,A})_i$, so that $(V_i)_{x:\,A} = (V_{x:\,A})_i$. Therefore, on the one hand: $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_1)_{x:\,A}]$, which implies $\mathfrak{A} \models^+_{\mathsf{G}} \forall x_{/Y}\varphi[V_1]$; and on the other hand: $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_2)_{x:\,A}]$ which implies $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_2]$ (by lemma 6.2.5). Moreover, it is easily verified that the V_i are Z-saturated in V, and that $V = V_1 \cup V_2$. Thus: $\mathfrak{A} \models^+_{\mathsf{G}} (\forall x_{/Y}\varphi \vee_{/Z} \psi)[V]$.

$(d', -)$ $\mathfrak{A} \models^-_{\mathsf{G}} (\forall x_{/Y}\varphi \vee_{/Z} \psi)[V]$ iff $\mathfrak{A} \models^-_{\mathsf{G}} \forall x_{/Y}\varphi[V]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V]$, iff $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[V_{x:\,f}]$ for some Y-independent $f : V \to A$, and $\mathfrak{A} \models^-_{\mathsf{G}} \psi_{/x}[V_{x:\,f}]$ (by lemma 6.2.5). Equivalently, $\mathfrak{A} \models^-_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,f}]$, which is in turn equivalent to: $\mathfrak{A} \models^-_{\mathsf{G}} \forall x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})[V]$

◁

Extraction of quantifiers over conjunctions now follows easily:

Corollary 6.4.6 (Quantifier extraction over $\wedge$) *If $Qx_{/Y}\varphi \wedge_{/Z} \psi$ is semi-regular and if x does not occur in ψ, then:*

(c″) $(\forall x_{/Y}\varphi \wedge_{/Z} \psi) \equiv_S \forall x_{/Y}(\varphi \wedge_{/Zx} \psi_{/x})$

(d″) $(\exists x_{/Y}\varphi \wedge_{/Z} \psi) \equiv_S \exists x_{/Y}(\varphi \wedge_{/Zx} \psi_{/x})$

Proof: Apply De Morgan's laws to lemma 6.4.5 (c') and (d') respectively. ◁

We now prove two lemmas that show that in some cases adding the independence of x at the main connective isn't necessary. In particular, they show that a slash for x at the main connective can always be omitted if the main connective isn't slashed for other variables.

Lemma 6.4.7 (Omitting variables, part I) *If $\exists x_{/Y}(\varphi_1 \vee_{/Zx} \varphi_2)$ is a semi-regular IFG-formula, and if $Z \subseteq Y$, then*

$$\exists x_{/Y}(\varphi_1 \vee_{/Zx} \varphi_2) \equiv_S \exists x_{/Y}(\varphi_1 \vee_{/Z} \varphi_2)$$

In particular, it always holds that $\exists x_{/Y}(\varphi \vee_{/x} \varphi_2) \equiv_S \exists x_{/Y}(\varphi \vee \varphi_2)$.

Informally, we can explain this as follows. In the situation of the lemma, Eloïse gets to make two subsequent moves: the first for the quantification $\exists x_{/Y}$, the next for the disjunction $\vee_{/Z(x)}$. A choice function f for the quantification assigns a value to x on the basis of the values assigned to the variables in $X - Y$. If she has to disregard the value assigned to x through f when making her choice at the disjunction, she can recalculate this value on the basis of the values assigned to the variables in $X - Z$, because $X - Y \subseteq X - Z$ by $Z \subseteq Y$. So, it does not matter whether we add the x to the slash of the main disjunction or not.

Note that the fact that the second move is a disjunction move, is not relevant to this argument: it holds as well for subsequent existential quantifiers:

Lemma 6.4.8 *If $\exists x_{/Y}\exists y_{/Zx}\varphi$ is semi-regular, and if $Z \subseteq Y$, then:*

$$\exists x_{/Y}\exists y_{/Zx}\varphi \equiv_S \exists x_{/Y}\exists y_{/Z}\varphi.$$

Neither is it relevant that the two moves are subsequent: as long a no assignments to variables are overwritten (which can be guaranteed by safeness and *regularity*), the same informal argument holds for Eloïse's moves deeper in the formula.

This (informal) generalization sheds some light on the issue of the 'slashing convention', i.e. the assumption that existential quantifiers are always independent of each other. This assumption is one of the features of Hintikka's IF-logic ([Hin96], point (vii) on p. 63); see e.g. [Jan02] for a discussion of this convention. The general version of this lemma learns us for which formulas the evaluation could alter by imposing the convention: e.g. for the typical example $\forall z\exists x\exists y_{/z}[y = z]$, which does not satisfy the condition that $Z \subseteq Y$ (in this case: $Z = \{z\}$ while $Y = \emptyset$). (We could see the general version of this lemma as an instance of the game-theoretic *Thompson transformation* of Inflation-Deflation; see section 4.8.1 of this thesis.)

We now only prove the special case of lemma 6.4.7 formally, in terms of the inductive clauses for satisfaction:

Proof: (of lemma 6.4.7) Suppose that the left formula, $\exists x_{/Y}(\varphi_1 \vee_{/Zx} \varphi_2)$, is semi-regular. Because the right formula has the same free and bound variables, it follows that the left and right formula are S-compatible.

In all cases, let $\mathfrak{A}$ be a suitable model, and $V \subseteq A^X$ a safe set of valuations for the given formulas. Then $x \notin X$. Suppose $Z \subseteq Y$.

(+) We will prove for any Y-independent $f : V \to A$ that elements of $V_{x:\,f}$ coincide out of Zx if and only if they coincide out of Z. It follows then for any subset W of $V_{x:\,f}$: W is Zx-saturated in $V_{x:\,f}$ if and only if W is Z-saturated in $V_{x:\,f}$. From the satisfaction condition for disjunction, it easily follows that

$$\mathfrak{A} \models^+_{\mathsf{G}} (\varphi_1 \vee_{/Zx} \varphi_2)[V_{x:\,f}] \text{ if and only if } \mathfrak{A} \models^+_{\mathsf{G}} (\varphi_1 \vee_{/Z} \varphi_2)[V_{x:\,f}]$$

and hence (because this holds for any Y-independent f):

$$\mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}(\varphi_1 \vee_{/Zx} \varphi_2)[V] \text{ if and only if } \mathfrak{A} \models^+_{\mathsf{G}} \exists x_{/Y}(\varphi_1 \vee_{/Z} \varphi_2)[V]$$

So let $f : V \to A$ be Y-independent and $v, w \in V$. We show that $v(x\colon f(v))$ and $w(x\colon f(w)) \in V_{x:\,f}$ coincide out of Zx iff they coincide out of Z. From right to left, this immediate. From left to right: if $v(x\colon f(v)) \sim_{Zx} w(x\colon f(w))$, then $v \sim_Z w$ (we use that $x \notin X$), and hence $v \sim_Y w$ by the assumption that $Z \subseteq Y$. Then $f(v) = f(w)$, so: $v(x\colon f(v)) \sim_Z w(x\colon f(w))$.

(-) The equivalence of $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}(\varphi_1 \vee_{/Zx} \varphi_2)$ and $\mathfrak{A} \models^-_{\mathsf{G}} \exists x_{/Y}(\varphi_1 \vee_{/Z} \varphi_2)$ is immediate: if we write out the satisfaction conditions given by definition 5.3.3, the independence conditions for $\exists$ and $\vee$ disappear.

$\triangleleft$

If we extract a universal quantifier $\forall x_{/Y}$ over a disjunction $\vee_{/Z}$, the addition of x under the slash of the main disjunction is never necessary, provided that the 'right disjunct' (ψ in lemma 6.4.5) is slashed for x (cf. Example 6.4.2):

Lemma 6.4.9 (Omitting variables, part II) *If $\forall x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})$ is a semi-regular IFG-formula, and if x does not occur in ψ, then*

$$\forall x_{/Y}(\varphi \vee_{/Zx} \psi_{/x}) \equiv_S \forall x_{/Y}(\varphi \vee_{/Z} \psi_{/x})$$

Proof: Again, assuming that the left formula is semi-regular, it easily follows that the left and right formula are S-compatible. Let $\mathfrak{A}$ be a suitable model, and $V \subseteq A^X$ a safe set of valuations for the given formulas. Then $x \notin X$.

(+) We prove that $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,A}]$, is equivalent to $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Z} \psi_{/x})[V_{x:\,A}]$. From left to right this follows easily from the fact that Zx-saturation of a $W \subseteq V_{x:\,A}$ implies Z-saturation.

For the converse, assume we have $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Z} \psi_{/x})[V_{x:\,A}]$. Pick $W_1, W_2 \subseteq V_{x:\,A}$ that are Z-saturated in $V_{x:\,A}$ with $W_1 \cup W_2 = V_{x:\,A}$, such that $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[W_1]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[W_2]$. By lemma 6.2.1, we may take the W_i to be disjoint. From W_1, W_2, we will construct sets W_1', W_2' that are Zx-saturated in $V_{x:\,A}$ with $W_1' \cup W_2' = V_{x:\,A}$ and $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[W_1']$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[W_2']$.

First, let V_2 be the projection of W_2 to X, so: $V_2 := \{v \in V | v(x{:}\, a) \in W_2 \text{ for some } a \in A\}$. Then $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_2]$ by lemma 6.2.5, and again by lemma 6.2.5, this implies $\mathfrak{A} \models^+_{\mathsf{G}} \psi_{/x}[(V_2)_{x:\,A}]$. We prove that $(V_2)_{x:\,A}$ is Zx-saturated in $V_{x:\,A}$: if $v \in V_2, w \in V$, and $v(x{:}\, a) \sim_{Zx} w(x{:}\, b)$, then $v \sim_Z w$ (because $x \notin X$); pick $c \in A$ such that $v(x{:}\, c) \in W_2$ (such c exists because $v \in V_2$), then $w(x{:}\, c) \in W_2$ by Z-saturation of W_2 in $V_{x:\,A}$, and hence $w(x{:}\, b) \in (V_2)_{x:\,A}$.

Now let $V_1 := V - V_2$, then $(V_1)_{x:\,A} = V_{x:\,A} - (V_2)_{x:\,A}$, hence $V_{x:\,A} = (V_1)_{x:\,A} \cup (V_2)_{x:\,A}$, and because $W_2 \subseteq (V_2)_{x:\,A}$ we have: $(V_1)_{x:\,A} = V_{x:\,A} \setminus (V_2)_{x:\,A} \subseteq (V_{x:\,A} \setminus W_2) = W_1$ (for this last equality we use that the W_i were chosen to be disjoint). Therefore, $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_1)_{x:\,A}]$ by lemma 6.2.1. Moreover, $(V_1)_{x:\,A}$ is Zx-saturated in $V_{x:\,A}$ since it is the complement of the Zx-saturated $(V_2)_{x:\,A}$.

We have shown that the sets $W_1' := (V_1)_{x:\,A}$ and $W_2' := (V_2)_{x:\,A}$ satisfy our requirements. Hence: $\mathfrak{A} \models^+_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,A}]$

(-) Because the negative evaluation of a disjunction, the variables under the slash don't play a role, we have for any $f : V \to A$:

$$\mathfrak{A} \models^-_{\mathsf{G}} (\varphi \vee_{/Z} \psi_{/x})[V_{x:\,f}] \text{ if and only if } \mathfrak{A} \models^-_{\mathsf{G}} (\varphi \vee_{/Zx} \psi_{/x})[V_{x:\,f}].$$

It easily follows that $\mathfrak{A} \models^-_{\mathsf{G}} \forall x_{/Y}(\varphi \vee_{/Z} \psi_{/x})[V]$ iff $\mathfrak{A} \models^-_{\mathsf{G}} \forall x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})[V]$.

◁

Combining results, we can now state the following equivalences with respect to quantifier extraction:

Theorem 6.4.10 (Quantifier extraction over connectives) *Given an IFG-formula of the form $Qx_{/Y}\varphi \vee_{/Z} \psi$ that is semi-regular, and such that x does not occur in ψ, then:*

$(\exists, \vee)$ $\exists x_{/Y}\varphi \vee_{/Z} \psi \equiv_S \exists x_{/Y}(\varphi \vee_{/Zx} \psi_{/x})$
and, under the condition that $Z \subseteq Y$:
$\exists x_{/Y}\varphi \vee_{/Z} \psi \equiv_S \exists x_{/Y}(\varphi \vee_{/Z} \psi_{/x})$.

$(\forall, \vee)$ $\forall x_{/Y}\varphi \vee_{/Z} \psi \equiv_S \equiv_S \forall x_{/Y}(\varphi \vee_{/Z} \psi_{/x})$

In particular:

$$Qx_{/Y}\varphi \vee \psi \equiv_S Qx_{/Y}(\varphi \vee \psi_{/x})$$

The cases $(\forall, \wedge)$ and $(\exists, \wedge)$ correspond to the cases $(\exists, \vee)$ and $(\forall, \vee)$ respectively (by De Morgan's laws).

Corollary 6.4.11 *If θ' is an IFG-sentence obtained from a regular IFG-sentence θ by application of quantifier extraction to one of its subformulas, then $\theta \equiv_G \theta'$.*

Note that regularity of θ implies that no subformula contains a variable both free and bound (semi-regularity), but to apply quantifier extraction to a subformula of the form $Qx_{/Y}\varphi \vee_{/Z} \psi$, we still need to make sure that x does not occur (bound) in ψ. When we say 'application of quantifier extraction', we imply that this is precondition is satisfied.

Proof: By the previous theorem, the substitution principle 6.3.16 (here we need the regularity of θ) and by the fact that G-equivalence and S-equivalence coincide on sentences (lemma 6.3.9). ◁

6.5 Prenex normal form

In the previous sections, we have reconstructed all building blocks necessary to support a prenex normal form theorem for IFG-formulas. Among the improvements made to the original system of [CK99], which allowed for some unexpected and disturbing signaling phenomena, were a correction to the existential clause of the inductive notion of satisfaction (Definition 5.3.3) and the introduction of a new notion of equivalence (Definition 6.3.7). These allowed us to prove new versions of Renaming of bound variables (theorem 6.3.17), Substitution of equivalents (theorem 6.3.16), and Quantifier extraction (theorem 6.4.10) in terms of these new notions.

We have to be careful that these results require regularity of the formulas involved, and S-compatibility of all formulas as precondition for (the transitivity of) S-equivalence. Fortunately, in the chain of formulas we get in the proof of our prenex normal form theorem, all formulas will have the same free variables, while the sets of bound variables increase monotonously. This ensures that lemma 6.3.8 is applicable for transitivity.

Theorem 6.5.1 (Prenex Form) *Any regular IFG-formula θ is S-equivalent to an IFG-formula in prenex form, with the same free variables as θ.*

Proof: Even if θ is regular, the same variable can occur bound by two or more quantifiers (if they have mutually disjoint scopes). By repeated application of theorems 6.3.17 (Change of Bound Variables) and 6.3.16 (Substitution of Equivalents), we obtain a chain of equivalent formulas

$$\theta = \theta_0 \equiv_S \theta_1 \equiv_S \cdots \equiv_S \theta_k = \theta'$$

such in θ' no two quantifiers bind the same variable. Regularity is preserved, because for each $i \in \{0, \ldots, k-1\}$, $Fv(\theta_i) = Fv(\theta_{i+1})$ and $Bv(\theta_{i+1}) = Bv(\theta_i) \cup \{u_i\}$ for some variable $u_i \notin Fv(\theta_i) \cup Bv(\theta_i)$.

By repeated application of theorem 6.4.10 (Quantifier Extraction) and again theorem 6.3.16, we obtain a further chain of equivalent formulas

$$\theta' = \theta'_0 \equiv_S \theta'_1 \equiv_S \cdots \equiv_S \theta'_k = \theta''$$

all with the same free and bound variables, such that θ'' is in prenex form. By transitivity of $\equiv_S$ on the formulas involved: $\theta \equiv_S \theta''$. ◁

In the prenex normal form that we get from this theorem, slashed connectives may still occur. To eliminate these, and get a stronger normal form result, we need a result like Theorem 2.3 from [CK99]: elimination of 'imperfect information connectives'. In the next section, we will discuss this theorem, point out how this theorem suffers from similar problems as the ones encountered before, and we prove an alternative formulation for it.

6.6 Elimination of slashed connectives

In this section, we check whether it is possible to eliminate slashes that occur at connectives. In other words: we check if any IFG-formula is equivalent to an IFG-formula in which no slashed disjunctions (or conjunctions) occur. Theorem 2.3 in [CK99, p. 24], 'Elimination of imperfect information connectives', claims this is the case (for G-equivalence). The main part of this theorem amounts to the following:

Quote 6.6.1 (Theorem 2.3 [CK99, p. 24]) *If $\mathfrak{A}$ is a suitable model containing at least two elements, $V \subseteq A^X$ is safe for $\varphi \vee_{/Y} \psi$, and $s, t \notin X$, then*

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi \vee_{/Y} \psi[V]$$

iff

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \exists s_{/Y} \exists t_{/Y}[(s = t \wedge \varphi) \vee (s \neq t \wedge \psi)][V]$$

(By lemma 6.4.8, we may omit the s in the second quantifier $\exists t_{/Y,s}$ of the original formulation)

Again, we can give two kinds of counterexamples to this claim. As before, the introduction of the new quantifiers may block signals from 'outside' (signals for $\varphi \vee_{/Y} \psi$ through the variables s, t may be overwritten if we evaluate the formulas with respect to a set of valuations $V \subseteq A^X$ with $s, t \in X$). We do not give an explicit counterexample for this problem. Reformulating the general claim of Theorem 2.3 of [CK99] in terms of S-equivalence rather than G-equivalence, will solve this problem.

A counterexample of a second kind shows how the new quantifiers may introduce new signaling possibilities *inside* the formula:

Example 6.6.2: (Quantifiers vs connectives) Consider the formula

$$\exists x_{/u}[x = u] \vee_{/y} \exists x_{/u}[x = u]$$

(so, in terms of the theorem: $\varphi = \psi = \exists x_{/u}[x = u]$ and $Y = \{y\}$). Let $\mathfrak{A}$ be a model containing 3 elements, say $A = \{0, 1, 2\}$, and take $V = A^{\{u,y\}}$. It is straightforward to verify that

$$\mathfrak{A} \not\models^+_{\mathsf{G}} \exists x_{/u}[x = u] \vee_{/y} \exists x_{/u}[x = u][V]:$$

This situation is part of the evaluation of the sentence $\forall u \forall y(\exists x_{/u}[x = u] \vee_{/y} \exists x_{/u}[x = u])$ in $\mathfrak{A}$. This sentence is only true in models containing at most two elements: the two different values Eloïse can choose for x (at the quantification over x in φ, and at the quantification over x in ψ), are not enough to cover the three possible values Abélard can choose for u.

According to the claim above, the chosen formula should be equivalent to:

$$\exists s_{/y} \exists t_{/y}[(s = t \wedge \exists x_{/u}[x = u]) \vee (s \neq t \wedge \exists x_{/u}[x = u])]$$

But the existential quantifier(s) create new possibilities for Eloïse: in her first move, she can assign the value of u to s, then choose t equal to s, and satisfy φ (the left copy of $\exists x_{/u}[x = u]$) by choosing the value of s for x. Note that this strategy does not violate any of the independence conditions.

(A careful reader may have noticed that the fact that the main disjunction is slashed for y, plays no role of importance in this example. Indeed, taking $Y = \emptyset$ would have been a counterexample to the claim as well, but arguably a less convincing one, as there would be no slashed connectives to eliminate.) $\diamond$

This example shows a fundamental difference between quantifiers and connectives: connectives can make a binary case distinction, but cannot really signal *values*. Existential quantifiers *can* do so.

A way to avoid this discrepancy, would be to require the models to interpret two distinct constants, **0** and **1** say, by two distinct domain elements (thereby excluding one-element models). Instead of the double existential quantification, the case distinction can then be made by just one existential quantifier:

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi \vee_{/Y} \psi[V]$$

iff

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \exists s_{/Y}[(s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)][V]$$

Note that requiring one special constant would not be enough: consider the following variant of the example above (for simplicity, we now take $Y = \emptyset$)

$$\mathfrak{A} \models^{+}_{\mathsf{G}} \exists s[(s = \mathbf{0} \wedge \exists x_{/u}[x = u]) \vee (s \neq \mathbf{0} \wedge \exists x_{/u}[x = u])][V].$$

Eloïse could still choose the value of s equal to the value of u. At the disjunction, $V_{s:\, f} = \{us\colon 00, 11, 22\}$ can be split in $W_1 = \{us\colon 00\}$ and $W_2 = \{us\colon 11, 22\}$. In both cases, Eloïse wins by choosing the value of s for x.

It is interesting to see that even though Hintikka does not prove an elimination theorem like Theorem 2.3 of [CK99], the slashed connectives in IF-logic *are* eliminated in the translation procedure from IF-logic to Σ^1_1(see e.g. [Hin96], formula (3.16) on page 52). There, the constant $\mathbf{0}$ is used as special constant, and implicitly, it is assumed that the models contain at least two elements. If we translate the sentence

$$\forall u \forall y(\exists x_{/u}[x = u] \vee_{/y} \exists x_{/u}[x = u])$$

from our Example 6.6.2 above, according to the procedure described (loc. cit.), it suffers from the same difference between connective and quantifier demonstrated by the Example 6.6.2. Because of the already present assumption that models contain at least two elements, the problem can easily be avoided by assuming two distinct special constants.

The previous observations suggest the following alternative for the part of Theorem 2.3 of [CK99] as quoted in Quote 6.6.1:

Lemma 6.6.3 (Elimination of slashed connectives) *Let φ,ψ be IFG-formulas. Then for any suitable model $\mathfrak{A}$ with distinct interpretations for the constants $\mathbf{0}$ and $\mathbf{1}$, $V \subseteq A^X$ safe for $\varphi \vee_{/Y} \psi$, and $s \notin X$:*

$$\mathfrak{A} \models^{\pm}_{\mathsf{G}} \varphi \vee_{/Y} \psi[V]$$

iff

$$\mathfrak{A} \models^{+}_{\mathsf{G}} \exists s_{/Y}[(s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)][V].$$

Proof: Let $c_0, c_1 \in A$ be the interpretations of $\mathbf{0}$ and $\mathbf{1}$ in $\mathfrak{A}$ respectively. We divide the proof in three cases:

$(+, \Rightarrow)$ First, suppose $\mathfrak{A} \models^{+}_{\mathsf{G}} \varphi \vee_{/Y} \psi[V]$. Determine disjoint $V_1, V_2 \subseteq V$ such that $V = V_1 \cup V_2$, V_1, V_2 Y-saturated in V and $\mathfrak{A} \models^{+}_{\mathsf{G}} \varphi[V_1]$ and $\mathfrak{A} \models^{+}_{\mathsf{G}} \psi[V_2]$. Let $f : V \to A$ be the characteristic function of V_2:

$$f(v) = \begin{cases} c_0 & \text{if } v \in V_1 \\ c_1 & \text{if } v \in V_2 \end{cases}$$

Then f is Y-independent by Y-saturation in V of V_2 and V_1. Furthermore, because $(V_i)_{s:\,f}$ is of the form $V_i \times \{c\}^{\{s\}}$ (where c is c_0 for V_1 and c_1 for V_2), we can use lemma 6.2.2 to show that

$$\mathfrak{A} \models^+_{\mathsf{G}} \varphi[(V_1)_{s:\,f}] \text{ and } \mathfrak{A} \models^+_{\mathsf{G}} \psi[(V_2)_{s:\,f}].$$

It follows easily that

$$\mathfrak{A} \models^+_{\mathsf{G}} (s = \mathbf{0} \wedge \varphi)[(V_1)_{s:\,f}] \text{ and } \mathfrak{A} \models^+_{\mathsf{G}} (s = \mathbf{1} \wedge \psi)[(V_2)_{s:\,f}],$$

and thus: $\mathfrak{A} \models^+_{\mathsf{G}} (s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)[V_{s:\,f}]$. By definition, it follows that: $\mathfrak{A} \models^+_{\mathsf{G}} \exists s_{/Y}[(s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)][V]$.

$(+, \Leftarrow)$ Conversely, assume that $\mathfrak{A} \models^+_{\mathsf{G}} \exists s_{/Y}[(s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)][V]$. Let $f : V \to A$ be independent of Y such that $\mathfrak{A} \models^+_{\mathsf{G}} (s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{0} \wedge \psi)[V_{s:\,f}]$.
Determine $W_1, W_2 \subseteq V_{s:\,f}$ such that $W_1 \cup W_2 = V_{s:\,f}$, and

$$\mathfrak{A} \models^+_{\mathsf{G}} (s = \mathbf{0} \wedge \varphi)[W_1] \text{ and } \mathfrak{A} \models^+_{\mathsf{G}} (s = \mathbf{1} \wedge \psi)[W_2].$$

For any $v \in V$, we write v_s for the valuation $v(s\colon f(v)) \in V_{s:\,f}$. Because $v_s \in W_1$ iff v_s satisfies $s = \mathbf{0}$, we have $W_1 = \{v_s \in V_{s:\,f} | v_s(s) = f(v) = c_0\}$. Similarly: $W_2 = \{v_s \in V_{s:\,f} | f(v) = c_1\}$.
Now define $V_1 := \{v \in V | f(v) = c_0\}$ and $V_2 := \{v \in V | f(v) = c_1\}$. By the fact that f is independent of Y, the V_i are Y-saturated in V. Furthermore, $W_1 \cup W_2 = V_{s:\,f}$ implies that $f(v) \in \{c_0, c_1\}$ for all $v \in V$, and hence: $V_1 \cup V_2 = V$.
The one issue left to check is whether $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[V_1]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V_2]$. For φ this holds by the fact that $\mathfrak{A} \models^+_{\mathsf{G}} \varphi[W_1]$, that $W_1 = V_1 \times \{c_0\}^{\{s\}}$, and by lemma 6.2.2. Similarly for ψ and V_2.
This verifies that $\mathfrak{A} \models^+_{\mathsf{G}} \varphi \vee_{/Y} \psi[V]$.

$(-, \Leftrightarrow)$ We will use following s-variation of V:

$$\begin{aligned} V_{s:\,0} &:= \{v(s/c_0) | v \in V\} &= V \times \{c_0\}^{\{s\}}; \\ V_{s:\,1} &:= \{v(s/c_1) | v \in V\} &= V \times \{c_1\}^{\{s\}}. \end{aligned}$$

The following are then equivalent:

- $\mathfrak{A} \models^-_{\mathsf{G}} \varphi \vee_{/Y} \psi[V]$
- $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[V]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V]$ (by definition)
- $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[V_{s:\,0}]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V_{s:\,1}]$ (by lemma 6.2.2)
- $(\mathfrak{A} \models^-_{\mathsf{G}} (s = \mathbf{0})[V_{s:\,A} \setminus V_{s:\,0}]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \varphi[V_{s:\,0}])$ and
$(\mathfrak{A} \models^-_{\mathsf{G}} (s = \mathbf{1})[V_{s:\,A} \setminus V_{s:\,1}]$ and $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V_{s:\,1}])$
- $\mathfrak{A} \models^-_{\mathsf{G}} (s = \mathbf{0} \wedge \varphi)[V_{s:\,A}]$ and $\mathfrak{A} \models^-_{\mathsf{G}} (s = \mathbf{1} \wedge \psi)[V_{s:\,A}]$ (by definition)
- $\mathfrak{A} \models^-_{\mathsf{G}} (s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)[V_{s:\,A}]$ (by definition)

- $\mathfrak{A} \models_{\mathsf{G}}^{-} \exists s_{/Y}[(s = \mathbf{0} \wedge \varphi) \vee (s = \mathbf{1} \wedge \psi)[V]$ (by definition)

$\triangleleft$

This result, however straightforward, is not a satisfactory replacement for Theorem 2.3 of [CK99]. In fact, it only works for models that have distinct interpretations for the constants $\mathbf{0}, \mathbf{1}$, and thereby excludes one-element models. This restriction is similar to the restriction to *IF-safe* models which was necessary to ensure soundness of the translation procedures from IF-logic to Σ_1^1 and back, which, by the way, also eliminates slashed disjunctions (cf. section3.2 of this thesis).

The following alternative is more complex in terms of the number of extra quantifiers needed, but solves all problems with Theorem 2.3 of [CK99] without imposing extra requirements on the models:

Theorem 6.6.4 (Elimination of slashed connectives) *Let s, t and u be distinct variables that don't occur in the IFG-formula $\theta = \varphi \vee_{/Y} \psi$, and let $Z \subseteq Fv(\theta)$ be the set of all free variables occurring under the slashes in φ. If θ is semi-regular, then:*

$$(\varphi \vee_{/Y} \psi) \equiv_S (\forall s \forall t[s = t] \wedge (\varphi \vee \psi)) \vee$$
$$(\exists s \exists t[s \neq t] \wedge \forall s_{/Z} \forall t_{/Z}[s = t \vee \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)]])$$

Proof: Denote the right formula in the equivalence by θ'. Note that θ, θ' are S-compatible by the requirement that no variables occur both free and bound in θ, and the requirement that the extra bound variables s, t, u in θ' don't occur in θ. Let $\mathfrak{A}$ be a suitable model for θ, and let $V \subseteq A^X$ be a safe set of valuations for θ and θ'.

We must prove:

$$\mathfrak{A} \models_{\mathsf{G}}^{\pm} \theta[V] \text{ iff } \mathfrak{A} \models_{\mathsf{G}}^{\pm} \theta'[V] \tag{6.3}$$

We distinguish two main cases:

1. $|A| = 1$ (i.e. $A = \{a\}$), or
2. $|A| \geq 2$

In the first case, we either have $V = \emptyset$ –in which case the equivalence (6.3) is trivial by first remark made directly after definition 5.3.3– or V is the singleton set $\{v\}$, where v is the valuation with $v(x) = a$ for all $x \in X$. (Note that for singleton sets of valuations, all independence conditions are vacuous.)

For positive satisfaction: suppose $\mathfrak{A} \models_{\mathsf{G}}^{+} (\varphi \vee_{/Y} \psi)[V]$. Because the independence condition for the disjunction is vacuous on V, it follows that $\mathfrak{A} \models_{\mathsf{G}}^{+} (\varphi \vee \psi)[V]$. Furthermore, it's easy to see that $\mathfrak{A} \models_{\mathsf{G}}^{+} \forall s \forall t(s = t)[V]$, so the first main disjunct of θ' is satisfied in $\mathfrak{A}$ by V, hence: $\mathfrak{A} \models_{\mathsf{G}}^{+} \theta'[V]$. The converse holds because $\mathfrak{A}$ and V can impossibly satisfy $\exists s \exists t(s \neq t)$, so, from $\mathfrak{A} \models_{\mathsf{G}}^{+} \theta'[V]$, it follows that $\mathfrak{A}$ and V must satisfy the first main disjunct of θ'. This implies $\mathfrak{A} \models_{\mathsf{G}}^{+} (\varphi \vee \psi)[V]$, which implies $\mathfrak{A} \models_{\mathsf{G}}^{+} (\varphi \vee_{/Y} \psi)[V]$ because the independence conditions are vacuous for V.

For negative satisfaction: suppose $\mathfrak{A} \models^-_{\mathrm{G}} (\varphi \vee_{/Y} \psi)[V]$. Then also $\mathfrak{A} \models^-_{\mathrm{G}} \varphi \vee \psi[V]$, so $\mathfrak{A}$ and V satisfy the left disjunct of θ' negatively. Furthermore, $\mathfrak{A} \models^-_{\mathrm{G}} \exists s \exists t(s \neq t)[V]$, so $\mathfrak{A}$ and V also satisfy the right disjunct of θ' negatively. Hence: $\mathfrak{A} \models^-_{\mathrm{G}} \theta'[V]$. The converse holds, because it follows from $\mathfrak{A} \models^-_{\mathrm{G}} \theta'[V]$ that $\mathfrak{A} \models^-_{\mathrm{G}} (\varphi \vee \psi)[V]$, and hence $\mathfrak{A} \models^+_{\mathrm{G}} (\varphi \vee_{/Y} \psi)[V]$ (because the independence conditions play no role in the negative clause for satisfaction for disjunction).

Now for the case that $|A| \geq 2$. It is easy to see that then: $\mathfrak{A} \not\models^+_{\mathrm{G}} \forall s \forall t(s = t)[V]$ and $\mathfrak{A} \models^+_{\mathrm{G}} \exists s \exists t(s \neq t)[V]$.

For positive satisfaction, from left to right: suppose $\mathfrak{A} \models^+_{\mathrm{G}} \varphi \vee_{/Y} \psi$. Determine disjoint $V_1, V_2 \subseteq V$ that are Y-saturated in V such that $V = V_1 \cup V_2$, $\mathfrak{A} \models^+_{\mathrm{G}} \varphi[V_1]$ and $\mathfrak{A} \models^+_{\mathrm{G}} \psi[V_2]$. Because V is safe for θ', we can define sets of valuations $W_1, W_2 \subseteq V_{st\colon A}$ by $W_1 := V \times \{(st\colon aa) | a \in A\}, W_2 := \{(st\colon ab) | a, b \in A, a \neq b\}$. Let $f : W_2 \to A$ be defined by

$$f(v(st\colon ab)) := \begin{cases} a & \text{if } v \in V_1 \\ b & \text{if } v \in V_2 \end{cases}$$

Then f is independent of Y by the fact that the V_i are Y-saturated, $(W_2)_{u\colon f} = V_1 \times \{(stu\colon aba) | a, b \in A, a \neq b\} \cup V_2 \times \{(stu\colon abb) | a, b \in A, a \neq b\}$, while by lemma 6.2.2:

$$\mathfrak{A} \models^+_{\mathrm{G}} \varphi[V_1 \times \{(stu\colon aba) | a, b \in A, a \neq b\}], \text{ and:}$$

$$\mathfrak{A} \models^+_{\mathrm{G}} \psi[V_2 \times \{(stu\colon abb) | a, b \in A, a \neq b\}]$$

Hence, respectively:

- $\mathfrak{A} \models^+_{\mathrm{G}} (u = s \wedge \varphi) \vee (u = t \wedge \psi)[(W_2)_{u\colon f}]$,
- $\mathfrak{A} \models^+_{\mathrm{G}} \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)][W_2]$,
- $\mathfrak{A} \models^+_{\mathrm{G}} (s = t) \vee \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)][V_{st\colon A}]$ (because $\mathfrak{A} \models^+_{\mathrm{G}} (s = t)[W_1]$ and $W_1 \cup W_2 = V_{st\colon A}$),
- $\mathfrak{A} \models^+_{\mathrm{G}} \forall s_{/Z} \forall t_{/Z}[(s = t) \vee \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)]][V]$ (the independence of Z of the universal quantifiers plays no role in the positive evaluation)
- $\mathfrak{A} \models^+_{\mathrm{G}} \exists s \exists t(s \neq t) \wedge \forall s_{/Z} \forall t_{/Z}[(s = t) \vee \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)]][V]$
- $\mathfrak{A} \models^+_{\mathrm{G}} \theta'[V]$

Conversely, suppose $\mathfrak{A} \models^+_{\mathrm{G}} \theta'[V]$. Then respectively:

- $\mathfrak{A} \models^+_{\mathrm{G}} s = t \vee \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)][V_{st\colon A}]$ (because no valuations in V satisfy the left disjunct of θ')
- $\mathfrak{A} \models^+_{\mathrm{G}} \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)][V \times \{(st\colon ab) | a, b \in A, a \neq b\}]$

- $\mathfrak{A} \models^+_{\mathrm{G}} \exists u_{/Y}[(u = s \wedge \varphi) \vee (u = t \wedge \psi)][V \times \{(st\colon a_0 b_0)\}]$
  if we fix any $a_0, b_0 \in A$ with $a_0 \neq b_0$, by lemma 6.2.1
- $\mathfrak{A} \models^+_{\mathrm{G}} (u = s \wedge \varphi) \vee (u = t \wedge \psi)[(V \times \{(st\colon a_0 b_0)\})_{u\colon f}]$
  for some $f : V \times \{(st\colon a_0 b_0)\} \to A$ that is Y-independent
- $\mathfrak{A} \models^+_{\mathrm{G}} (u = s \wedge \varphi)[(V_1 \times \{(st\colon a_0 b_0)\})_{u\colon f}]$ and $\mathfrak{A} \models^+_{\mathrm{G}} (u = t \wedge \psi)[(V_2 \times \{(st\colon a_0 b_0)\})_{u\colon f}]$ for some $V_1, V_2 \subseteq V$ with $V_1 \cup V_2 = V$. Then necessarily $f(v) = a_0$ iff $v \in V_1$, and $f(v) = b_0$ iff $v \in V_2$ (so the V_i are automatically disjoint, and Y-saturated by Y-independence of f). Hence:
- $\mathfrak{A} \models^+_{\mathrm{G}} \varphi[V_1 \times \{(stu\colon a_0 b_0 a_0\}]$ and $\mathfrak{A} \models^+_{\mathrm{G}} \varphi[V_2 \times \{(stu\colon a_0 b_0 b_0\}]$. Then:
- $\mathfrak{A} \models^+_{\mathrm{G}} \varphi[V_1]$, $\mathfrak{A} \models^+_{\mathrm{G}} \psi[V_2]$ by lemma 6.2.2. Because the V_i are Y-saturated:
- $\mathfrak{A} \models^+_{\mathrm{G}} \varphi \vee_{/Y} \psi[V]$

Hence: $\mathfrak{A} \models^+_{\mathrm{G}} \theta[V]$ iff $\mathfrak{A} \models^+_{\mathrm{G}} \theta'[V]$.

For negative satisfaction, it is enough to prove that the following are equivalent:

$$\mathfrak{A} \models^+_{\mathrm{G}} (\neg\varphi \wedge \neg\psi)[V] \tag{6.4}$$

$$\mathfrak{A} \models^+_{\mathrm{G}} \exists s_{/Z} \exists t_{/Z}[s \neq t \wedge \forall u[(u \neq s \vee \neg\varphi) \wedge (u \neq t \vee \neg\psi)]][V] \tag{6.5}$$

Assume (6.4). Fix $a, b \in A$ with $a \neq b$, then $\mathfrak{A} \models^+_{\mathrm{G}} (\neg\varphi \wedge \neg\psi)[V_{u\colon A} \times \{(st\colon ab)\}]$ by lemma 6.2.2. It then follows that:

$$\mathfrak{A} \models^+_{\mathrm{G}} (u \neq s \vee \neg\varphi) \wedge (u \neq t \vee \neg\psi)[V_{u\colon A} \times \{(st\colon ab)\}]$$

(take the empty set of valuations for the left disjuncts), and:

$$\mathfrak{A} \models^+_{\mathrm{G}} s \neq t \wedge \forall u[(u \neq s \vee \neg\varphi) \wedge (u \neq t \vee \neg\psi)][V \times \{(st\colon ab)\}].$$

Using two constant (hence Z-independent) functions that assign a and b respectively to all valuations, (6.5) follows easily.

Conversely, assume (6.5). Then

$$\mathfrak{A} \models^+_{\mathrm{G}} s \neq t \wedge \forall u[(u \neq s \vee \neg\varphi) \wedge (u \neq t \vee \neg\psi)][(V_{s\colon f})_{t\colon g}]$$

for some $f : V \to A$ and $g : V_{s\colon f} \to A$ that are both Z-independent. Necessarily, $f(v) \neq g(v(s\colon f(v)))$ for all $v \in V$. Also, both:

$$\mathfrak{A} \models^+_{\mathrm{G}} (u \neq s \vee \neg\varphi)[((V_{s\colon f})_{t\colon g})_{u\colon A}]$$

$$\mathfrak{A} \models^+_{\mathrm{G}} (u \neq t \vee \neg\psi)[((V_{s\colon f})_{t\colon g})_{u\colon A}]$$

By lemma 6.2.1, it follows that

$$\mathfrak{A} \models^+_{\mathrm{G}} (u \neq s \vee \neg\varphi)[((V_{s\colon f})_{t\colon g})_{u\colon f^*}]$$

$$\mathfrak{A} \models^+_{\mathrm{G}} (u \neq t \vee \neg\psi)[((V_{s\colon f})_{t\colon g})_{u\colon g^*}]$$

where $f^*, g^* : (V_{s:\,f})_{t:\,g} \to A$ are defined by $f^*(w) = w(s)$ and $g^*(w) = w(t)$ respectively. Then all valuations in $((V_{s:\,f})_{t:\,g})_{u:\,f^*}$ satisfy $u = s$, hence necessarily:

$$\mathfrak{A} \models^+_{\mathsf{G}} \neg\varphi[((V_{s:\,f})_{t:\,g})_{u:\,f^*}]$$

and similarly

$$\mathfrak{A} \models^+_{\mathsf{G}} \neg\psi[((V_{s:\,f})_{t:\,g})_{u:\,g^*}]$$

Note that because $s, t \notin Z \subseteq X$, the functions f^*, g^* are Z-independent, as were f, g. By lemma 6.2.6, it follows that $\mathfrak{A} \models^+_{\mathsf{G}} \neg\varphi[V]$ and $\mathfrak{A} \models^+_{\mathsf{G}} \neg\psi[V]$, hence (6.4).

So, we also have: $\mathfrak{A} \models^-_{\mathsf{G}} \theta[V]$ iff $\mathfrak{A} \models^-_{\mathsf{G}} \theta'[V]$. ◁

With this result, we now have:

Corollary 6.6.5 *Every regular IFG-formula is S-equivalent to a regular IFG-formula without slashed connectives.*

Proof: This follows from repeated application of theorem 6.6.4, the substitution principle theorem 6.3.16 (this requires regularity), following a similar chain of arguments to those in the prenex normal form theorem 6.5.1. ◁

Corollary 6.6.6 *Every regular IFG-formula is S-equivalent to a regular IFG-formula without slashed connectives which is in prenex form.*

Proof: First apply the previous corollary, then the prenex normal form theorem 6.5.1. It follows from the Quantifier Extraction theorem 6.4.10 that if a formula contains no slashed connectives before quantifier extraction, then neither does the result of the quantifier extraction. ◁

Corollary 6.6.7 *Every regular IFG-sentence is G-equivalent to a regular IFG-sentence without slashed connectives which is in prenex form.*

Proof: By lemma 6.3.9 and the previous corollary. ◁

With this last result, we may conclude that the main result from [CK99] holds indeed for (the restricted class of) *regular IFG-sentences*, despite all the nuances we had to introduce in the steps that lead to it.

6.7 Conclusions

In this chapter we have proved some general properties of satisfaction for IFG-formulas, and several equivalence schemes for IFG-formulas, leading to a prenex normal form theorem for *regular* IFG-*sentences*. In this normal form, no slashed connectives occur.

In comparison with results for IF-logic, these results are symmetric in their treatment of truth and falsity, while both language and semantics are designed with as little implicit rules as possible on the application and interpretation of the slash operator (e.g. no *slashing convention*).

We have demonstrated several signaling phenomena related to situations that from a classical viewpoint seem unproblematic. They occur if a variable occurs in some sense *double*: both free and bound in one formula, in a nested quantification, or both bound in the formula and in the domain of the set of valuations with which the formula is evaluated. We showed how these situations invalidate many results claimed in [CK99]. It motivated us to introduce the notions of S-equivalence and regularity, which prove to enable reformulations and proofs of the invalidated claims of [CK99].

We think that similar precautions are needed to prove equivalence schemes using Hodges' trump semantics, although formulating them will be easier by the fact that the domains of the valuations are a fixed parameter with the formula (and not an external parameter).

Chapter 7

General conclusions and open issues

In this final chapter, we look back on the work of the previous chapters and highlight some conclusions. We also list a number of open issues as inspiration for future research.

7.1 Hintikka's approach

We started this thesis by giving a general introduction to the basic ideas behind Independence Friendly logic (IF-logic) and game theoretical semantics as they were presented and advocated by Hintikka, in particular in his book *The Principles of Mathematics Revisited* ([Hin96]).

In chapter three, we gave a formal account of the Skolemization procedure used by Hintikka to obtain Σ_1^1-truth conditions for IF-sentences. These Σ_1^1-truth conditions, and the fact that every Σ_1^1-sentence may be viewed as a truth condition for some IF-sentence, make it easy to infer properties for IF-logic from the properties of Σ_1^1. In fact, we feel that the Skolemization procedure (based on the usual Skolemization procedures for first order logic and partially ordered quantification) constitutes Hintikka's basic interpretation of IF-sentences, and not the game theoretic concepts of game theoretical semantics.

There are two kinds of observations that support this.

First, we have shown that Hintikka's approach to IF-logic focuses on truth only (recognizably so in the definition of the syntax and the notion of equivalence he uses). This corresponds to the fact that Skolemization delivers truth-conditions only. A game theoretic approach would have included falsity as second dimension. It treats both players equally in the definition of the semantic games, and truth and falsity correspond to the separate questions under which conditions each of them has a winning strategy. That in semantic games for IF-sentences these are *truly* separate questions is demonstrated by the result of Burgess [Bur03]. To ignore falsity is a non-trivial choice from a game-theoretic perspective.

Second, a game theoretic analysis of some formulas (most notably Hodges' example (3.10)), gives different evaluations than the Σ_1^1-truth condition obtained by Hintikka's Skolemization procedure. From a game-theoretic perspective, we would need a Skolemization procedure that also takes previous existential quantifiers into account (while the traditional procedure only takes universal quantifications into account). But Hintikka, probably when realizing this, did not repair the Skolemization procedure. Instead, he introduces a game-theoretically counterintuitive provision, *the slashing convention*, in order for the Skolemization procedure to remain unchanged.

Taking the Skolemization procedure as primitive, it is still possible to also include falsity aspects. For this, we generalized the original IF-syntax to its symmetric extension $\mathcal{L}_{IFS}$, which allowed us to obtain the falsity condition of an IF-sentence as the truth condition of the negation of the sentence.

In the rest of the thesis, we have left what we called 'Hintikka's approach'. We did not take the Skolemization procedure as primitive (and did not adopt the slashing convention), and defined strategies for *both* players in terms of the semantic games (and not by a syntactic procedure). We switched to the IFG-language, which is built up from first order atomic formulas rather than first order sentences in negation normal form; it generalizes the IF-language in the following respects: it includes *open* formulas, negation may occur in front of any subformula, and any variable may occur under a slash of any quantifier or connective.

7.2 Game theory as framework for logic

In chapter 4 we have taken the 'game theoretical' in game theoretical semantics seriously, and actually formalized semantic games for IFG-sentences into the game-theoretic standard model of games in extensive form. In the game-theoretic model, strategies in semantic games can be taken to be sets of choice functions. These choice functions in turn, can be viewed as Skolem functions of the type that also takes existential quantifications into account.

Furthermore, we noted that the structure of the game trees for semantic games was rather specific, and so is the structure of the *information sets*, the component modeling the imperfect information induced by the slash operator. We investigated to which extent allowing connectives under the slash would extend the class of game trees corresponding to semantic games, but only in a restricted number of cases did we find a sensible interpretation.

Having a game-theoretic formalization of semantic games, we can study game-theoretic notions and results in order to find their counterparts for IFG-logic. From this perspective, we studied the Thompson transformations. These allowed us to formulate some equivalence schemes for IFG-sentences, but they also showed us that –without the possibility of putting connectives under the slash, and with the specific structure of the semantic games– the applicability of the transformations is restricted. Therefore, the main result associated with the transformations, viz. that they can bring any game in extensive form into *reduced normal form*, does not directly transfer to IFG-logic. It was interesting to note that games of *imperfect recall*, which are usually avoided in game theory, are omnipresent in semantic games for IF-logic (in almost all cases, by the slashing convention) and in IFG-logic.

In the line of looking for relations between logical and game-theoretical notions and results, there are many open questions and issues. We list some of them below (when we say 'semantic games', we mean semantic games *for IFG-sentences*):

- Semantic games have a specific structure, so not every arbitrary game in extensive form is a semantic game. For example, a game resulting from the application of the Thompson transformation *coalescence* to a semantic game, is no longer a semantic game. But the result is *equivalent* to a semantic game. We wonder: is every arbitrary game equivalent (by Thompson or other transformations) to a semantic game? In other words: would there be a procedure that transforms an arbitrary game into a game that has the structure of a semantic game?

- Related to the previous question: could there exist a logic (logical language with game semantics) which is characterized by the Thompson transformations as axioms? (Question suggested by Wilfrid Hodges in relation to [Dec04].)

- Already mentioned in correlation to the Thompson transformations is the game-theoretic *reduced normal form*. On the logical side we have several normal forms, which we could try to relate to game-theoretic ones: negation normal form (which is not noticeable game-theoretically, cf. section 4.5), prenex normal form (which could be seen as the result of Thompson transformations), the Skolem normal form (which we automatically get for IF-sentences after application of the translation into Σ_1^1 and back, cf. [Hin96, p. 60]; it seems related to the reduced normal form although it only preserves *truth*, or in other words: Eloïse's powers), or the distributive normal form (cf. [Vää02], this also only preserves truth).

- There seems to be an essential difference between quantifiers and connectives if it comes to independence. Independence of the value of a certain variable can easily be expressed in terms of (sets of) valuations, but our discussion of section 4.6 shows that it is harder to grasp what independence of a connective would mean in general. Maybe we could explain the difference between the two types of logical constants (or: two types of moves?) in game theoretic terms. We could attempt (as suggested by Eric Pacuit) to use the game-theoretic distinction between *imperfect* information and *incomplete* information (where the latter means uncertainty about some part of the game *structure*).

- Games (in extensive form) are in a natural manner related to processes (cf. [vB02a]). The game theoretic approach to logic with the independence operator may lead to applications in the theory of parallel processing, as demonstrated by [Bra00].

7.3 The slash operator and its interpretations

With the previous two sections in mind, let's look at the two main interpretations of the slash operator in IF(G)-logic. One could say that in Hintikka's approach, the interpretation of the slash operator is principally in terms of *independence*, while, in the game-theoretic context, the slash operator is interpreted by *imperfect information*. In this thesis we have encountered some examples for IF(G)-sentences that have (in some sense) surprising evaluations.

An example that incorporates several elements of 'surprise' is Hodges' formula $\forall x \exists z \exists y_{/x}[x = y]$. This formula evaluates differently under Hintikka's approach (making it *true only in one-element models*) than under a game theoretic interpretation (making it *logically valid*).

As noted in section 3.9, the **dependence relation** of the quantifiers in this formula is **not transitive**: while y may depend on z and z may depend on x, y is not allowed to depend on x. (This example thereby shows that the linear notation with the slash operator can express more dependence patterns than partially *ordered* quantification.) Even though a relation (as mathematical notion) need not be transitive, the term *dependence* (in general) usually implies transitivity. We therefore get conceptual conflicts when we try to interpret this formula.

Hintikka solves the counterintuitive situation by introducing the slashing convention. At least, this resolves the specific intransitivity of dependence in Hodges' example by making y independent of z as well. But in game-theoretic terms, this convention is very counterintuitive: why would Eloïse forget *all* her own moves, while she may remember previous ones of the opponent. The slashing convention introduces **imperfect recall** into many formulas. On the other hand, in section 4.7 we noted that, even without the slashing convention, the semantic games for Hodges' formula are of imperfect recall (because Eloïse forgets knowledge she previously had, viz. the value of x). Eloïse overcomes this however by a 'trick' we call *signaling*. Whether or not this trick can be considered as *cheating*, depends on how we take the interpretation of the concept of imperfect recall. In our formalization of the semantic games, the trick corresponds to a legal winning strategy.

Note that the slashing convention does not exclude all types of intransitivity of the dependence relation between quantifiers: in $\forall x \forall z \exists y_{/x}\psi$ the dependence relation is just as intransitive as in Hodges' example. But this intransitivity causes no problems, because Eloïse has no control over the value of z, so she can no longer *use* the intransitivity for signaling.

Furthermore, while the slashing convention solves the problem of *unwanted* intransitivity of the dependence relation, it does not get rid of conceptual problems surrounding the interpretation of independence: with the slashing convention, the IF-formula $\exists x \exists y[x = y]$ is still logically valid even while y is not allowed to depend on x.

These examples make us ask fundamental and difficult questions like what the notion of dependence of quantifiers really means. Valuable work in this respect has been done in [Jan02], which works towards a formalization of *independence* of quantifiers (and choices) that is in accordance with our intuitions.

7.4 Semantics for open formulas

After the first chapters, in which only sentences were evaluated, we introduced the semantics of [CK99] for *open* IFG-formulas in chapter 5. We presented the semantics by giving a game theoretic definition first, followed by an inductive definition using the same parameters. The game theoretic notion of satisfaction allows for a good comparison with the game theoretical semantics defined by Hintikka (for sentences only), and the inductive notion is useful for proving equivalence schemes without using translations to other logical systems (as we demonstrate in chapter 6). We proved that these two notions of satisfaction coincide, and also used the inductive version to prove explicitly that IFG-logic with this semantics extends classical first order logic.

A closer study of the following issues would help us to get a better understanding of the semantics for open IFG-formulas:

- In a certain sense, the sets of valuations code a *context* for an open formula that is evaluated. In the (counter-)examples we gave in chapters 5 and 6, we could always find a sentence φ such that the open formula ψ and set of valuations V of the examples could be seen as part of the (positive) inductive evaluation of φ. A natural question to ask is: given a model $\mathfrak{A}$ and a formula ψ, can we construct such a sentence for any arbitrary set of valuations $V \subseteq A^X$?

  If the answer is 'yes' (which we would expect on the basis of our experience), such a sentence need not be unique (e.g. if $\psi := \exists y_{/x}[x = y]$, $\mathfrak{A} := \langle\{0,1\},\ldots\rangle$ and $V := \{xz\colon 00, 11\}$, then φ could be $\forall x \exists z \exists y_{/x}[x = y]$, but also: $\forall z \exists x \exists y_{/x}[x = y]$).

  Note also that if we find such φ, it only holds for positive evaluation: $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V]$ if and only if $\mathfrak{A} \models^+_{\mathsf{G}} \varphi$. We do not yet have that $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V]$ if and only if $\mathfrak{A} \models^-_{\mathsf{G}} \varphi$, because the inductive evaluation of $\mathfrak{A} \models^-_{\mathsf{G}} \varphi$ goes along different clauses, hence gives a different sets of valuations. (For example, if we take ψ as above and $\varphi := \forall x \exists z \exists y_{/x}[x = y]$, the second step of the evaluation now gives $V' = \{xz\colon c0, c1\}$ for some $c \in \{0,1\}$, hence $V' \neq V$. A sentence φ' with $\mathfrak{A} \models^-_{\mathsf{G}} \varphi'$ iff $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V]$ would be $\varphi' := \exists x \forall z \exists y_{/x}[x = y]$.)

- Another interesting question is that of *satisfiability* of a given IFG-formula ψ and suitable model $\mathfrak{A}$ (due to Peter van Emde Boas): can we find a set of valuations V for ψ in $\mathfrak{A}$ such that $\mathfrak{A} \models^+_{\mathsf{G}} \psi[V]$ (and similarly for $\mathfrak{A} \models^-_{\mathsf{G}} \psi[V]$)? If so, what is the complexity of finding such V?

- The main difference of the semantics of [CK99] with Hodges' Trump semantics (cf. [Hod97a]) is that in Hodges' system, formulas come with a fixed domain for the sets of variables with which they are evaluated. This makes a formula $\psi(x)$ different from the formula $\psi(x,z)$, even if z does not occur in ψ. These domains restrict the class of the sentences in which they can occur: $\psi(x)$ can only occur under the scope of precisely one quantification (over x), while $\psi(x,z)$ can only occur in the scope of precisely two quantifications (over x and over z). This does not seem to be a big difference, but we are careful after the subtle results and warnings of chapter 6, which also involve the domains of the valuations. It would be worth wile to compare our semantics with Trump semantics in more detail.

7.5 The requirement of regularity

Hilbert and Ackermann defined the language of first order predicate logic in such a way that a variable never occurs both free and bound within a single formula, nor can a formula contain nested quantification over the same variable [HA59, p. 74]. In this thesis, we have called formulas that satisfy these properties *regular* (cf. definition 6.3.15). Although first order languages are nowadays usually defined without these restrictions (as we also did in our chapter 2), imposing them does not affect the properties of the logic: by renaming of bound variables, any irregular first order formula may be transformed into a regular equivalent.

In chapters 5 and 6 however, we have showed that the difference between regular and irregular formulas does have semantical consequences for both IF- and IFG-logic. First, we showed that a restriction to regular formulas is crucial in Hintikka's claim that IF-logic is a conservative extension of classical first order logic (cf. section 5.6).

For IFG-logic, the matter is even more subtle, as is shown by the efforts in chapter 6 in order to free the prenex normal form theorem of [CK99] from flaws related to irregular formulas. Complicating factor in this logic is the fact that we also have to be careful with the domains of the valuations we evaluate a formula with: these act just like free variables of the formula (even if they do not actually occur free in the formula), and should therefore be included in the restriction to regularity (the domain of the valuations may not contain variables that occur bound in the formula). Regularity for IFG-logic with the semantics we presented in this thesis is thereby a notion with aspects outside the syntax of the formulas. In this respect, we see a big advantage in Hodges' choice to let the domains be a parameter that comes with the formula.

The problems with irregularity can be seen as an artefact caused by coding the positions in the semantic games by *valuations*. If in a later move a new value is assigned to a variable, the old value in the valuation is overwritten and therefore becomes inaccessible to the players. In this way, nested quantification over the same variable, or free variables that occur bound later in the formula, introduce another form of imperfect information in an intrinsic way, that may become significant in combination with the slash operator. Or, in Hintikka's approach: with the Skolemization procedure. To avoid this, we must follow the good example of Hilbert and Ackermann, and restrict ourselves to regular situations.

7.6 Some possible directions for future research

We highlight a couple of possible directions for further research, which we think are interesting to pursue.

We have not paid much attention to (non-)constructivistic aspects of the definitions and results in this thesis. Chapter 10 of [Hin96] is dedicated to a **constructivist approach** to game theoretical semantics and IF-logic (as in Hintikka's approach). The idea is basically to restrict the interpretation of the Skolem functions in the Σ_1^1-truth conditions to *recursive* functions. Of course, we would also have to interpret the connectives and predicates constructively. We note that in [CS98] it is argued this idea to turn to recursive Σ_1^1, will not deliver an *intuitionistic* approach, as it would conflict with Brouwer's thesis that every real-valued function is continuous. These are all subtle issues, and they deserve to be studied in more detail.

A non-constructivistic element that is *incorporated* in game theoretic semantics, is the axiom of choice. This is a consequence of the conception of strategies as (consisting of) functions, as is usual in game theory. Alternatively, one could define a different concept of strategy, viz. in terms of a *relation*, and **avoid incorporating the axiom of choice** in the logic (as we noted on page 96). We have not looked extensively into the consequences of this switch, but the proofs of some results in chapter 6 use the function approach (e.g. lemmas 6.2.5, 6.4.7). We can therefore not yet say whether, or in which form, the prenex normal form theorem would hold for this conception of strategy (if we keep avoiding the axiom of choice, of course).

An interesting attempt to get a grip on the relation between the notion of *winning strategy* in semantic games, and **proof** for *first order* validity, is presented in [Bny04]. It uses the idea, also present in game semantics for linear logic ([Abr02]), that (one of the) players may choose to return to a previous position in the game and play again from there. In linear logic, a similar idea is used to define one of the implications. Maybe similar ideas may also lead to some game-theoretic notion of **implication** for IF(G)-logic.

In section 3.8, we collected some reflections on game-theoretic negation, and in section 3.6, we mentioned the possibility of extending IF-logic to **EIF-logic**, by adding a weak, sentence-initial, contradictory negation. It would be interesting to look at possible connections between EIF-logic and other systems with strong and weak negations, most notably **Nelson's logic of constructible falsity** ([Nel49], or sections 3.1-3.2 of [Jas94]). In this constructive propositional logic, a formula is true if *we have a proof.* The strong negation of a formula is true if *we have a* *refutation*, while the weak negation of it is satisfied if *we have no proof.* Replace 'proof' by 'strategy for Eloïse', and 'refutation' by 'strategy for Abélard', and we see how this corresponds to the strong and weak negation of EIF-logic. (However, when trying to elaborate this connection, through studying the algebras associated with Nelson's logic (cf. [MV01]), we hit the problem that in EIF-logic the weak negations only occur sentence-initially, while the strong negations –by negation normal form– occur only at atomic level.)

After our study of the proposal of [Hin96], we have become a bit skeptical about Hintikka's claim that IF-logic is truly *new and revolutionary*, and even whether IF-logic is truly *first order.* This is because in Hintikka's approach, IF-logic almost coincides with Σ_1^1: *classical* existential *second order* logic. Nevertheless, the ideas and concepts that are presented in the book have proved to be intriguing, inspiring, and puzzling. They force us to be more precise about 'old' notions because translating them to IF(G)-logic may not be as straightforward as we expected. In this respect, the book is also an inspiration for the philosophy of logic and mathematics.

Bibliography

[Abr02] Samson Abramsky. Lecture notes on game semantics, 2002. Course notes for the course "Interactive Models of Logic and Computation", as part of the North American Summer School on Logic, Language and Information, Stanford 2002.

[Bar85] Jon Barwise. Model-theoretic logics: Background and aims. In Jon Barwise and Solomon Feferman, editors, *Model-Theoretic Logics*, pages 3–23. Springer-Verlag, 1985.

[vB00a] Johan van Benthem. 'Hintikka Self-Applied': an essay on the epistemic logic of imperfect information games, January 2000. Preliminary version of [vB05].

[vB00b] Johan van Benthem. Logic in games, Autumn 2000. Lecture notes for the Master of Logic program at the ILLC, University of Amsterdam. Draft version, manuscript versions 2001, 2002. Electronically available at `http://staff.science.uva.nl/~johan/` .

[vB01] Johan van Benthem. Games in dynamic epistemic logic. *Bulletin of Economic Research*, 53(4):219–248, 2001.

[vB02a] Johan van Benthem. Extensive games as process models. *Journal of Logic, Language and Information*, 11(3):289–313, 2002.

[vB02b] Johan van Benthem. Logic games are complete for game logics. *Studia Logica*, 75(2):183–203, 2002.

[vB03] Johan van Benthem. Logic and game theory: close encounters of the third kind. In G. Mints and R. Muskens, editors, *Games, Logic and Constructive sets*, volume 161, pages 3–22. Stanford: CSLI publications, 2003.

[vB04] Johan van Benthem. Probabilistic Features in Logic Games. In D. Kolak and J. Symons, editors, *Quantifiers, Questions and Quantum Physics.* Springer Verlag, New York, 2004. Invited lecture, Open Court Symposium in Honor of Jaakko Hintikka, APA Chicago.

[vB05] Johan van Benthem. The epistemic logic behind IF-games. In R.E. Auxier and L.E. Hahn, editors, *The Philosophy of Jaakko Hintikka*, volume 30 of *Library of Living Philosophers*. Southern Illinois University, 2005. Forthcoming.

[Bno04] Giacomo Bonanno. Memory and perfect recall in extensive games. *Games and economic behavior*, 47:237–256, 2004.

[Bny04] Denis Bonnay. Preuves et jeux sémantiques. *Philosophia Scientiae*, 8(2), November 2004. Special issue: "Logique et théorie des jeux", co-edited by Manuel Rebuschi and Tero Tulenheimo.

[Bra00] Julian Bradfield. Independence: logic and concurrency. In P.G. Clote and H. Schwichtenberg, editors, *Proceedings of the 14th International Workshop on Computer Science Logic (CSL 2000)*, number LNCS 1862 in Lecture Notes in Computer Science. Springer, August 2000.

[Bur03] John Burgess. A remark on Henkin sentences and their contraries. *Notre Dame Journal of Formal Logic*, 44(3), 2003.

[CDJ] Xavier Caicedo, Francien Dechesne, and Theo M.V. Janssen. Equivalence and quantifier rules in logics with imperfect information. (in preparation).

[CK99] Xavier Caicedo and Michal Krynicki. Quantifiers for reasoning with imperfect information and Σ_1^1-logic. In Walter A. Carnielli and Itala M. L. Ottaviano, editors, *Contemporary Mathematics*, volume 235, pages 17–31. American Mathematical Society, 1999.

[CS98] Roy Cook and Stewart Shapiro. Hintikka's revolution: The principles of mathematics revisited. *Journal of Logic, Language and Information*, 49:309–316, 1998.

[Cor98] David Corfield. Come the revolution. *Philosophical books*, 39(3):150–156, July 1998.

[vD91] Eric van Damme. *Stability and Perfection of Nash Equilibria*. Springer-Verlag, second, revised edition, 1991.

[Dec02] Francien Dechesne. An extensive game model for IF-logic. In *Proceedings of the NASSLLI02 Student Session*, Stanford, CSLI, 2002.

[Dec04] Francien Dechesne. Thompson transformations for IF-logic? In Peter McBurney et al., editor, *Proceedings of the workshop Knowledge and Games*. Department of Computer Science, University of Liverpool, July 2004.

[Dec05] Francien Dechesne. Falsity conditions for IF-sentences. *Philosophia Scientiae*, Spring 2005. Special issue with contributions to the conference *Philosophical Insights into Logic and Mathematics*, Nancy 2002, forthcoming.

[ER94] Susan Elmes and Philip J. Reny. On the strategic equivalence of extensive form games. *Journal of Economic Theory*, 62:1–23, 1994.

[End70] H. B. Enderton. Finite partially ordered quantifiers. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 16:393–397, 1970.

[Fit96] Melvin Fitting. *First order logic and automated theorem proving.* Graduate texts in computer science. Springer Verlag, Berlin, second edition, 1996.

[Fre79] Gottlob Frege. Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens, 1879. Reprinted in [vH67, p. 1–82] (English translation).

[GS53] David Gale and Frank Stewart. Infinite games with perfect information. *Ann. Math. Studies*, 28:245–266, 1953.

[Göd31] Kurt Gödel. Über Formal Unentscheidbare Sätze der Principia Mathematica und Verwandter Systeme. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931. Reprinted in English as [Göd92].

[Göd92] Kurt Gödel. *On Formally Undecidable Propositions of Principia Mathematica and Related Systems.* Dover Publications, New York, 1992.

[Gol98] Laurence Goldstein. Review of The Principles of Mathematics Revisited. *Philosophical Investigations*, pages 285–289, 1998.

[Hen61] Leon Henkin. Some remarks on infinitely long formulas. In *Infinitistic methods. Proceedings of the Symposium on Foundations of Mathematics, Warsaw.*, pages 167–183. Pergamon Press, 1961.

[HA59] David Hilbert and Wilhelm Ackermann. *Grundzüge der theoretischen Logik.* Springer Verlag, fourth edition, 1959.

[vH67] Jean van Heijenoort. *From Frege to Gödel – a source book in Mathematical Logic, 1879–1931.* Harvard University Press, 1967.

[Hin96] Jaakko Hintikka. *The Principles of Mathematics Revisited.* Cambridge University Press, 1996.

[Hin02a] Jaakko Hintikka. Hyperclassical logic (a.k.a. IF-logic) and its implications for logical theory. *The Bulletin of Symbolic Logic*, 8(3):404–423, September 2002.

[Hin02b] Jaakko Hintikka. Quantum logic as a fragment of Independence Friendly Logic. *Journal of Philosophical Logic*, 31:197–209, 2002.

[Hds98] Harold Hodes. Review of The Principles of Mathematics Revisited. *Journal of Symbolic Logic*, 63(4):1615–1623, 1998.

[Hod97a] Wilfrid Hodges. Compositional semantics for a language of imperfect information. *Bulletin of the IGPL*, 5 (4):539–563, 1997.

[Hod97b] Wilfrid Hodges. Review of The Principles of Mathematics Revisited. *Journal of Logic, Language and Information*, 6(4):457–460, 1997.

[Hod97c] Wilfrid Hodges. Some strange quantifiers. In Jan Mycielski, Grzegorz Rozenberg, and Arto Salomaa, editors, *Structures in Logic and Computer Science*, number 1261 in Lecture Notes in Computer Science, pages 51–65. Springer, 1997.

[Hod01a] Wilfrid Hodges. Dialogue Foundations: a sceptical look. *Supp Proc Arist Soc*, 75(1):17–32, 2001.

[Hod01b] Wilfrid Hodges. Formal features of compositionality. *Journal of Logic, Language and Information*, 10:7–28, 2001.

[Jan97] Theo M. V. Janssen. A compositional semantics for the game-theoretical interpretation of logic. In Paul Dekker, Martin Stokhof, and Yde Venema, editors, *Proceedings of the Eleventh Amsterdam Colloquium*, pages 181–185, December 1997.

[Jan02] Theo M. V. Janssen. Independent choices and the interpretation of IF-logic. *Journal of Logic, Language and Information*, 11:367–387, 2002.

[JD05] Theo M. V. Janssen and Francien Dechesne. Signaling in IF-games: a tricky business. In J. van Benthem et al., editor, *The age of alternative logics*. Kluwer, 2005. Forthcoming.

[Jas94] Jan Jaspars. *Calculi for Constructive Communication*. PhD thesis, University of Tilburg, 1994.

[Kre97] Philippe Kreutz. Review of The Principles of Mathematics Revisited. *Revue Internationale de Philosophie*, 2(100):288–290, 1997.

[Kuh53] Harold W. Kuhn. Extensive games and the problem of information. In H.W. Kuhn and A.W. Tucker, editors, *Contributions to the Theory of Games II*, number 28 in Annals of Mathematics Studies, pages 193–216. Princeton, 1953. Reprinted in [Kuh97, p. 46–68].

[Kuh97] Harold W. Kuhn, editor. *Classics in Game Theory*. Princeton University Press, 1997.

[KV89] Michal Krynicki and Jouko Väänänen. Henkin and function quantifiers. *Annals of Pure and Applied Logic*, 43:273–292, 1989.

[Lei94] Daniel Leivant. Higher order logic. In Dov M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Deduction Methodologies*, volume 2 of *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 229–299. Oxford Science Publications, 1994.

[Man96] Maria Manzano. *Extensions of first order logic.* Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1996.

[MV01] Luiz F. Monteiro and Ignacio Viglizzo. Construction of Nelson Algebras, 2001. Presented at XIV Coloquio LatinoAmericano de Álgebra, received through personal communication.

[Nel49] David Nelson. Constructible falsity. *Journal of Symbolic Logic*, 14(1):16–26, March 1949.

[NM47] John von Neumann and Oscar Morgenstern. *Theory of Games and Economic Behaviour.* Princeton University Press, 1947.

[OR94] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory.* MIT Press, 1994.

[PV03] Rohit Parikh and Jouko Väänänen. Finite information logic, 2003. Available at `http://www.sci.brooklyn.cuny.edu/cis/parikh/`.

[PR97] Michele Piccione and Ariel Rubinstein. On the interpretation of decision problems with imperfect recall. *Games and economic behavior*, 20:3–24, 1997.

[vR86] Arnoud van Rooij. *Analyse voor beginners.* Number 6 in Epsilon reeks. Epsilon Uitgaven, 1986.

[Rus03] Bertrand Russell. *The Principles of Mathematics.* Cambridge University Press, 1903.

[SV92] Gabriel Sandu and Jouko Väänänen. Partially ordered connectives. *Math. Logik Grundlagen Math.*, 38:361–372, 1992.

[San97] Gabriel Sandu. The logic of informational independence and finite models. *Logic Journal of the IGPL*, 5(1):79–95, 1997.

[San98] Gabriel Sandu. IF-logic and truth definition. *Journal of Philosophical Logic*, 27:143–164, 1998.

[San01] Gabriel Sandu. Signalling in languages with imperfect information. *Synthese*, 127(1-2):21–34, April-May 2001.

[SH01] Gabriel Sandu and Jaakko Hintikka. Aspects of compositionality. *Journal of Logic, Language and Information*, 10:49–61, 2001.

[SP01] Gabriel Sandu and Ahti Pietarinen. Partiality and games: Propositional logic. *Logic Journal of the IGPL*, 9(1):101–121, 2001.

[SP03] Gabriel Sandu and Ahti Pietarinen. Informationally independent connectives. In G. Mints and R. Muskens, editors, *Logic, Language and Computation*, volume 9, pages 23–41. Stanford: CSLI publications, 2003.

[Sch67] Joseph R. Schoenfield. *Mathematical Logic.* Addison-Wesley series in Logic. Addison Wesley, 1967.

[Sel75] Reinhard Selten. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975. Reprinted in [Kuh97, p. 317–354].

[Sev04] Merlijn Sevenster. Renaissance in IF-logic. In *Logic, Games and Philosophy: Foundational Perspectives, Abstracts*, pages 40–42, 2004. Abstract for the Prague International Colloquium, Sep. 28–Oct. 1.

[Ste92] Ian Stewart. *The Problems of Mathematics.* Oxford University Press, New York, 1992. new edition.

[SW01] Ulrich Schwalbe and Paul Walker. Zermelo and the early history of game theory. *Games and economic behavior*, 34(1):123–137, 2001. Contains an English translation of [Zer13].

[dS93] Harrie de Swart. *Logic: Mathematics, Language, and Philosophy*, volume I. Peter Lang, New York, 1993.

[dSVB99] Harrie de Swart, Tom Verhoef, and Renske Brands. Hintikka's "The Principles of Mathematics Revisited". *Logique et Analyse*, 159:281–289, 1999.

[Tar33] Alfred Tarski. The concept of truth in the languages of the deductive sciences (polish). *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matemat yczno-Fizycznych*, 34, 1933. Expanded English translation in [Tar83], pages 152-278.

[Tar83] Alfred Tarski. *Logic, Semantics, Metamathematics, papers from 1923 to 1938.* Hackett Publishing Company, Indianapolis, 1983.

[Ten98] Neil Tennant. Games some people would have all of us play. *Philosophia Mathematica*, 6:90–115, 1998.

[Tho52] F.B. Thompson. Equivalence of games in extensive form. Technical Report RM-759, The Rand Corporation, 1952. Reprinted in [Kuh97, p. 36–45].

[Tul02] Tero Tulenheimo. On IF-modal logic and its expressive power. In *Advances in Modal Logic*, 2002. Toulouse, Sep. 30–Oct. 2.

[Tul03] Tero Tulenheimo. *Independence-friendly modal logic: studies in its expressive power and theoretical relevance.* PhD thesis, University of Helsinki, 2003.

[Vää01] Jouko Väänänen. Second order logic and the foundations of mathematics. *Bulletin of Symbolic Logic*, 7:504–520, 2001.

[Vää02] Jouko Väänänen. On the semantics of informational independence. *Logic Journal of the IGPL*, 10(3):339–352, 2002.

[Wal70] Wilbur John Walkoe. Finite partially-ordered quantification. *Journal of Symbolic Logic*, 35(4):535–555, December 1970.

[WR13] Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*, volume I–III. Cambridge University Press, 1910,1912,1913.

[Zer13] Ernst Zermelo. Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels. In *Proc. Fifth Congress Mathematicians*, pages 501–504. Cambridge University Press, 1913.

Appendix A

About Abélard and Eloïse

Throughout this thesis, Abélard and Eloïse feature as the players of the abstract semantic games. The use of the names of Abélard (in the role of falsifier) and Eloïse (in the role of verifier), is common in the literature on logic games, mainly because of their mnemonic value (Abélard for the $\forall$-quantifier and Eloïse for the $\exists$-quantifier). In this appendix, for those who wonder who Abélard and Eloïse were, we give some historic background to these names.

The 'real' Abélard (Pierre Abélard, 1079–1142) and Eloïse (or Heloise, 1101–1162) are nowadays mostly remembered for their legendary love affair, which became literary history by the letters they wrote to each other while they lived in separation. Much of what we know about Abélard's life and his character, is conveyed to us in Abélard's *Historia Calamitatum*, which could be described as an autobiography written in the form of a letter.

Pierre Abélard came from a noble Breton family. It was usual for young men of his standing to opt for a military career. But instead, Abélard chose to study *dialectic*, i.e. philosophy, which at the time consisted mainly of the logic of Aristotle as transmitted through latin channels. He was very bright and had an unmistakable talent for teaching. At the early age of 22, he set up a very successful school, which soon rivalled the school of his former teacher, William of Champaux. Shortly after he started attending the theology lectures of St. Anselm at Laon, he was allowed to teach there as well. Around the year 1115, his fame reached a climax, and his lectures drew large numbers of students. He obtained the chair at Notre-Dame cathedral in Paris, and was nominated canon (this is a clergyman belonging to the chapter or staff of a cathedral or collegiate church).

At Notre-Dame, his eye was caught by the very young Eloïse, who lived there as the protégéé of her uncle Fulbert, canon at the Notre Dame. Not only was she beautiful, she also had an unusually broad knowledge. He seduced her while being her tutor. He saw no reason to be secretive about their relationship, but in fact there was at least one: her uncle. When Fulbert found out about it, he separated Abélard and Eloïse. But they kept meeting in secret; she got pregnant and gave birth to their son, Astrolabius. Abélard wanted to take his responsibility, and proposed to marry in an attempt not to offend Fulbert. At the same time, he wanted the marriage to remain a secret in order to keep his prospects for a career in the church open.

Even though Eloïse tried to convince him to keep his independence, they did marry. However, Fulbert didn't keep it a secret. Abélard helped Eloïse escape from her uncle and humiliation, to the convent of Argenteuil, and their son was sent to live with Eloïse's sister. Fulbert, thinking that Abélard had helped Eloïse just to be rid of her, took revenge by castrating him. While Abélard's ambitions to become reach a high position in the church, Eloïse became a nun.

In 1120, Abélard turned back to his studies, and started teaching anew, again with great success. But his views kept making him new enemies: for the rest of his life he was followed by persecution, trials, condemnations for his rationalistic doctrines. In 1121, having escaped from incarceration at the Abbey of St. Medard, he set up a school which he named the "Paraclete" (an appellaton of the Holy Ghost meaning *comforter*).

A few years later, he was absolved and reinstated as a monk. He made Eloïse Abess of the Paraclete, of which he still was the spiritual director. This gave him some opportunity to see her every once in a while. The reading of his *Historia Calamitatum*, moved her to write her first letter, followed by others, in which she step by step reconciles herself with the fact that they should live as brother and sister rather than as lovers.

In 1141, Abélard died in Cluny, while on his way to Rome, where he was to defend himself against a number of charges of herecy. His remains were secretly transferred to Paraclete, and given over to the loving care of Eloïse. When she died in 1164, she was laid to rest beside him. In 1817, they were transferred to the cemetery of Père Lachaise in Paris, where they still lay side-by-side in a monument erected for them out of materials from the Paraclete.

The correspondence of Abélard and Eloïse inspired the English poet Alexander Pope (1688–1744) to capture their love affair in his poem *Eloisa to Abelard.* The phrase "Eternal sunshine of a spotless mind", taken from this poem, was used for the title of a recent movie in which a company called Lacuna Inc. sells the opportunity to erase unwanted memories. (In a sense, they sell imperfect recall.)

(This brief description of the lives of Pierre Abélard and Eloïse is based on the entries for Abélard and Eloïse at the internet encyclopedia Wikipedia. The picture of Abélard and Eloïse used for the cover of this thesis comes from a 15th century *lettrine.*)

Samenvatting

In zijn boek *The Principles of Mathematics Revisited* (1996) betoogt de Finse filosoof en logicus Jaakko Hintikka dat de als 'klassiek' te boek staande eerste orde logica aan vervanging toe is. Hij stelt voor om de semantiek van Tarski te vervangen door een speltheoretische semantiek (Game Theoretical Semantics), en een operator '/' aan de logische taal toe te voegen, waarmee kwantoren (en voegtekens) onafhankelijk gemaakt kunnen worden van andere kwantoren: in de formule $\forall x \exists y R(x,y)$ mag de waarde van y afhangen van de waarde van x, in de formule $\forall x \exists y_{/x} R(x,y)$ is dit expliciet verboden. Het resultaat heet "Independence Friendly"-logica, afgekort: IF-logica.

Of een (gesloten) formule waar of onwaar is in een gegeven model, wordt bepaald door het bestaan van een winnende strategie voor een van de spelers (resp. Eloïse en Abélard) in een abstract evaluatiespel, dat we een *semantisch spel* noemen. De onafhankelijkheidsoperator in de logische taal maakt de semantische spelen tot spelen van onvolledige informatie, en deze zijn in het algemeen niet beslist (in de zin dat een van de spelers een winnende strategie heeft). Een opvallende eigenschap van IF-logica is dan ook dat de *wet van het uitgesloten derde* niet geldt.

In bovengenoemd boek beschrijft Hintikka de IF-logica en zijn eigenschappen zonder precieze formele definities of bewijzen te geven. Dit proefschrift: *Game, Set, Maths: een formele verkenning van logica met onvolledige informatie*, beoogt basisbegrippen van de IF-logica en haar semantiek op een mathematisch precieze wijze te geven, en de daarbij aan het licht komende subtiliteiten te bespreken.

In hoofdstuk 1 geven we een informele inleiding in het onderwerp en een kort overzicht van verwante literatuur. Hoofdstuk 2 bevat de notaties en basisbegrippen zoals deze gebruikt worden in dit proefschrift.

In hoofdstuk 3 volgen we Hintikka's benadering van IF-logica door de strategieën in semantische spelen gelijk te stellen aan rijtjes Skolem-functies. We schrijven Hintikka's op eerste-orde Skolemisatie gebaseerde vertaalprocedure van IF-logica naar Σ^1_1 tot in detail uit. (Σ^1_1 is eerste orde logica uitgebreid met tweede orde kwantificatie over functievariabelen.) De volgorde van de Skolemisatie-stappen bepaalt of eerdere existentiële kwantoren wel of niet opgenomen worden als argument in de Skolemfuncties. We laten zien dat verschillende volgordes voor IF-logica niet-equivalente uitkomsten opleveren (in tegenstelling tot voor eerste orde logica). Dit wordt gedemonstreerd door een belangrijk voorbeeld van Wilfrid Hodges, waarin de afhankelijkheidsrelatie van de kwantoren niet-transitief is. Om de voor klassieke logica gebruikelijke 'van buiten naar binnen'-volgorde te kunnen blijven hanteren voor IF-logica, heeft Hintikka een tegen-intuïtieve aanname nodig, namelijk dat de spelers hun eerdere zetten vergeten.

Hintikka's vertaalprocedure is gebaseerd op het gelijkstellen van de uitdrukkingskracht van een formule met zijn *waarheidsconditie.* Doordat de wet van het uitgesloten derde niet geldt, is *onwaarheid* echter een tweede dimensie in de uitdrukkingskracht van de formule. We laten zien hoe we *onwaarheidscondities* kunnen bepalen; hiervoor moeten we o.a. Hintikka's oorspronkelijke IF-taal eerst symmetrisch maken. We vertalen een resultaat van John Burgess voor Henkin-logica naar IF-logica; dit laat zien dat ieder paar van incompatibele Σ^1_1-zinnen te zien is als een paar van waarheidsconditie en onwaarheidsconditie voor een IF-zin.

Vanaf hoofdstuk 4 maken we ons losser van Hintikka's (eigen) werk en definities. We werken met een generalisatie van Hintikka's syntax, $\mathcal{L}_{IFG}$, gebaseerd op de taal uit [CK99].

We formaliseren de semantische spelen in het speltheoretische standaardmodel van de *extensive games.* Deze benadering van IF-logica ontbreekt in het werk van Hintikka zelf, maar is in recenter werk van anderen op dit gebied de standaard geworden. Dit levert een directe wiskundige formalisering op van het centrale concept van winnende strategie, op basis van een formalisering van het spel zoals het beschreven wordt door de regels (iets wat ontbreekt in Hintikka's benadering met Skolemfuncties als strategieën). Aan de hand van het model bespreken we in hoeverre we de taal zouden kunnen uitbreiden door niet alleen onafhankelijkheid van gekwantificeerde variabelen, maar ook van voegtekens toe te laten.

Op basis van de speltheoretische formalisering, verkennen we in hoeverre speltheoretische begrippen en resultaten kunnen bijdragen aan onze inzichten in de logica. We zien dat de in de speltheorie vermeden eigenschap van *onvolledige herinnering* in de semantische spelen voor IFG-zinnen eerder uitzondering dan regel is. Ook bespreken we de vier Thompson transformaties in relatie met logische equivalentieschema's voor IFG-logica.

Tot en met hoofdstuk 4 hebben we (zoals Hintikka zelf) alleen gewerkt met zinnen, d.w.z. *gesloten* formules. In de hoofdstukken 5 en 6 stappen we over op de semantiek van [CK99] om te kunnen werken met IFG-formules met vrije variabelen. Een open IFG-formule wordt geëvalueerd ten opzichte van een *verzameling* valuaties, om de onafhankelijkheidsoperator te kunnen interpreteren.

In hoofdstuk 5 signaleren we een probleem met de inductieve regel van deze semantiek voor de existentiële kwantor. Het probleem ontstaat als een zelfde variabele dubbel gebruikt wordt: dit kan zogeheten signalen blokkeren, en daarmee het verschil maken tussen het wel of niet bestaan van een winnende strategie. Met 'dubbel gebruik' bedoelen we geneste kwantificatie over dezelfde variabele, of het zowel vrij als gebonden voorkomen van een variabele. Uit de verschillende (tegen-)voorbeelden in de hoofdstukken 5 en 6, blijkt dat alle variabelen in het domein van de verzameling valuaties waarmee een IFG-formule geëvalueerd wordt, wat dit betreft gezien moeten worden als vrije variabelen van de formule (ook als ze niet vrij in de formule voorkomen). Situaties waarin dit 'dubbel gebruik' niet voorkomt, noemen we *regulier*. We laten zien dat beperking tot reguliere zinnen nodig is voor de claim dat Hintikka's IF-logic een conservatieve uitbreiding is van klassieke eerste orde logica.

In hoofdstuk 6 laten we zien dat ook de prenexnormaalvormstelling voor IFG-logica van [CK99] beperkt moet worden tot reguliere situaties: voor vrijwel alle lemma's en stellingen uit het artikel vinden we tegenvoorbeelden. We laten zien hoe deze gerepareerd kunnen worden. Het noodzaakt tot de invoering van een nieuw, minder strict, maar partieel equivalentie-begrip, met behulp waarvan de ondersteunende lemma's opnieuw bewezen worden.

Hoofdstuk 7 bevat een overzicht van conclusies en open vragen die binnen de beperkte termijn van dit onderzoek opgekomen maar niet opgelost zijn, en als inspiratie zouden kunnen dienen voor toekomstig werk.

Curriculum Vitae

Francien Dechesne werd op 26 oktober 1971 in Nijmegen geboren. Na het behalen van het gymnasium diploma aan het Dominicus College te Nijmegen, studeerde zij wiskunde aan de Katholieke Universiteit Nijmegen van 1989 tot 1996. Ook volgde zij vakken uit de bovenbouwstudie filosofie aan dezelfde universiteit. Ze studeerde *cum laude* af bij dr. Wim Veldman, op de scriptie: *Alles op een rijtje - Een intuïtionistische verkenning van totale ordeningen.* Fincancieel ondersteund door een VSB-beurs, bezocht zij van september 1996 tot maart 1997 een half jaar de vakgroep wiskunde van de University of California at Los Angeles (UCLA).

Vervolgens werkte zij als systeembeheerder van IBM-mainframes gedurende ruim drie jaar in dienst van automatiseringsbedrijf Pink Elephant te Zoetermeer. In september 2000 begon zij als assistent in opleiding aan het onderzoeksproject waar dit proefschrift het eindresultaat van vormt. Als promovenda van het Samenwerkingsorgaan Brabantse Universiteiten (SOBU), was zij zowel verbonden aan de faculteit Wijsbegeerte van de Universiteit van Tilburg, als aan de faculteit Wiskunde en Informatica van de Technische Universiteit Eindhoven. Tevens was zij lid van de landelijke Onderzoekschool Logica (OzsL).